AD No.

DDC FILE COPY

AD A049231

11

C.A.C.I.

WASHINGTON, D.C. OFFICES

11

JAN 31

F

MEASUREMENT OF UNIT EFFECTIVENESS
IN MARINE CORPS INFANTRY BATTALIONS

Final Technical Report

Sponsored by:
Defense Advanced Research Projects Agency

ARPA Order Number	3379
Program Code Number	T110
Program Element Number	62711E
Contractor	CACI, Inc.-Federal
Effective Date of Contract	March 9, 1977
Contract Expiration Date	October 31, 1977
Contract Number	MDA903-77-C-0199
Principal Investigator	Dr. Richard E. Hayes (703) 841-7800
Short Title of Work	Effectiveness of Infantry Battalions
Date of Report	October 31, 1977
Period Covered by Report	March 9, 1977 - October 31, 1977

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

CACI, INC.-FEDERAL: 1815 NORTH FORT MYER DRIVE, ARLINGTON, VIRGINIA 22209, TELEPHONE (703) 841-7800
A SUBSIDIARY OF CACI, INC. • WASHINGTON, D.C. • LOS ANGELES • HARRISBURG • NEW YORK CITY • THE HAGUE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 14 CAC3810002AG

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER: 9

4. TITLE (and Subtitle): 6 Measurement of Unit Effectiveness in Marine Corps Infantry Battalions. ~~Final Technical Report~~

5. TYPE OF REPORT & PERIOD COVERED: Final Report, 9 Ma[redacted]77-Octo[redacted]r 31, 1977

6. PERFORMING ORG. REPORT NUMBER: 9 Mar-31 Oct 77

7. AUTHOR(s): 10 Richard E. Hayes, Paul C. Davis, John J. Hayes, Farid Abolfathi, Bill Harvey

8. CONTRACT OR GRANT NUMBER(s): MDA903-77-C-0199, new 15

9. PERFORMING ORGANIZATION NAME AND ADDRESS: CACI, Inc.-Federal, 1815 North Fort Myer Drive, Arlington, Virginia 22209

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: ARPA Order [redacted]3379, Program Code No. TL10, Program Element Code 62711E

11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Advanced Research Projects Agency, 1400 Wilson Boulevard, Arlington, Virginia 22209

12. REPORT DATE: 11 31 Oct[redacted]r [redacted], [redacted]77

13. NUMBER OF PAGES: 148

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Readiness Branch, Operations and Training Division, Headquarters, U.S. Marine Corps, Washington, D.C. 20380

12 149p.

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Recommend unlimited distribution.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DDC
JAN 31 1978
F

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Evaluation, Mission Accomplishment, Combat Effectiveness, Performance, Historical Research, U.S. Marine Corps, Readiness, Battalion, Unit Attributes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The relative combat effectiveness of U.S. Marine Corps infantry battalions i 72 engagements spanning from World War II through Vietnam was estimated on the basis of the judgment of experienced Marine Corps officers about mission accomplishment. These same officers also identified the factors that they felt were crucial determinants of combat performance. Historical data were also collected on the units, the enemy forces involved, and the engagements. Multivariate statistics, including correlation, regression, and factor analysis, were used to examine the resulting data sets.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

407136 1B

20. Continued

The major findings of the analysis were that adaptive behavior by units -- reaction to the combat environment -- is the major discriminator between successful and unsuccessful performance. Three major types of activity appear to be related to mission accomplishment -- command and planning functions, which are strongly related to effective performance; supporting fires, which are moderately associated; and coordination functions, which are associated relatively weakly and appear to require effective command and planning before they make a difference. Analyses of over two dozen specific variables are presented in the report. A final section suggests innovations in the evaluation of exercise data to allow more effective measurement of the potential for combat effectiveness.

Volume I contains the substantive analyses including an executive summary, the research rationale, data collection efforts, statistical analyses, and suggestions for innovation. Volume II is a technical appendix that includes narrative descriptions of the 22 engagements, survey instruments, and data collection forms utilized by the research team.

A

PREFACE

The research team wishes to acknowledge the generous cooperation and support received from a number of individuals and organizations. The Cybernetics Technology Office of the Defense Advanced Research Projects Agency, which has a major interest in the area of combat effectiveness, sponsored these analyses. Lt. Col. Roy Gulick of that office, program manager for the effort, contributed valuable guidance and suggestions to the research project.

The Contracting Officer's Technical Representative, Col. Richard J. Johnson of the Evaluation Team, Readiness Branch, Operations Division of Headquarters Marine Corps, performed admirably in his role. His willingness and ability to act as effective liaison for the project with offices and organizations throughout the Washington, D.C., area and the Marine Corps were an invaluable asset. His prompt, thorough review of drafts, willingness to challenge, probe, and question, capability for sharing his experience and insights, and high standards of performance were of exceptional help to the effort. Through his good offices, the considerable talents of Col. Bettistore and the other members of the Readiness Branch were brought to bear on the project.

The research effort would have been nearly impossible if it had not been for the excellent professional assistance received from the Marine Corps Historical Center. Brig. Gen. E.H. Simmons, Director of the Center, took a personal interest in the project and ensured that the fine new research facilities there were available to the study team even before they were formally open to the public. Mr. Henry I. (Bud) Shaw was very generous with his time and energy. His first-hand knowledge of the available historical materials is tremendously impressive. His willingness to share his knowledge and insights

saved the research team a tremendous amount of time. Ms. Bonnet, the Archives Secretary, was an effective professional who tolerated massive intrusion into her limited space with cheerfulness and genuine goodwill.

Cooperation from the officers and commands within the Marine Corps was exceptional. A visit by the principal investigator to the base at Twentynine Palms, California, was particularly valuable. The personnel at the Marine Corps Air-Ground Combat Training Center (MCAGCTC) demonstrated effective professionalism, not only in the conduct of exercises, but also in their ability to explain, clearly and in detail, the rationale and procedures that they utilize. The hospitality of the officers and enlisted personnel in that command was genuine and is greatly appreciated.

A visit to the Naval War College at Newport, Rhode Island, was characterized by similar hospitality, interest, and helpfulness. The President of the College, Admiral Huntington Hardisty, and particularly Col. William Weise, USMC, Senior Marine Corps Representative on the faculty there, who took a personal interest in the project, were most gracious in allowing the use of their facilities for data collection.

In addition, a special word of thanks is due to all the Marine Corps officers, active duty and retired, who contributed their time and energy to the data coding. The professionalism they displayed added significantly to the content of the study. Long hours of concentration and hard work were cheerfully given by all concerned officer/students at the Naval War College, retired officers, officers awaiting school assignment, and volunteers from Headquarters Marine Corps. Without help from these people, the research could not have been completed. Their contributions have made possible whatever value the analyses produce. They are not, of course, responsible for problems or errors committed by the research team.

A number of different CACI staff members contributed to this project. The principal investigator was Dr. Richard E. Hayes, Manager of the

Policy Sciences Division. Military experience and expertise were provided by MG John J. Hayes, USA(Ret) and Dr. Paul Davis (Col, USA(Ret)). Construction of narratives was supported by Messrs. Gary Keynon and William Harvey. Help with questionnaire construction was provided by Drs. Bertram Spector and David McCormick. Mr. Harvey took primary responsibility for assembling the data set and initial selection of statistical approaches. Dr. Farid Abolfathi carried out most of the detailed, multivariate analyses.

The massive task of producing the paper necessary for the effort was handled with high professionalism by the Policy Sciences support staff led by Ms. Nancy Streeter and Ms. Donna Goodyear. The typing was cheerfully and quickly completed, at times at great price in sore eyes, by April Bailey, Lisa Dueno, Kathy Harris, and Jessica Johnson.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1. EXECUTIVE SUMMARY

BACKGROUND

One of the most crucial problems facing military commanders and defense decision-makers is the maintenance of an "effective" military capability. Creating combat-ready units in time of war, establishing standards and priorities for training, procuring, and staffing during periods of relative peace, and assuring a military capability that deters potential adversaries from dangerous adventures have long been central missions of the U.S. armed forces. Yet, measuring "military effectiveness" without data from combat experience remains a major unresolved research problem.

Military operations are complex undertakings involving a variety of social processes (such as C^3, planning, and so forth), physical movements, mechanical operations, and rapidly changing situations about which knowledge is always incomplete. These have traditionally been described in terms of functional areas -- command, operations, logistics, intelligence, and so forth. The failure of any one functional area <u>may</u> undermine the performance of a unit ("for want of a nail a shoe was lost..."). <u>But military effectiveness is more than the sum of its functional parts</u>. All functional aspects of an operation may go well, but a unit may still fail to accomplish its mission. Conversely, history is replete with examples of military units that achieved their missions despite functional failures.

The U.S. military has long recognized this and has therefore made commanders responsible not for mere technical proficiency within their units but for effective performance. Units that have fewer functional failures, overcome obstacles, and achieve their objectives are viewed as "successful." Focus is placed on outcomes -- the achievement of desired goals -- rather than the process involved in getting there. In short, military effectiveness can only be measured in terms of mission accomplishment. The research reported here was sponsored by the Cybernetics Technology Office of the Defense Advanced Research Projects Agency (ARPA/CTO) to determine whether a research methodology based on this principle could be developed and implemented.

OBJECTIVES

The research reported in this volume had three major objectives:

1. Improvement of understanding of the concept of combat effectiveness.
2. Development of algorithms (procedures) for projecting combat effectiveness.
3. Exploration of the implications of combat effectiveness for readiness.

As used in this report, the concept "combat effectiveness" refers to the ability of a unit to accomplish a military mission. As such, combat effectiveness refers to performance in a hostile environment, as distinguished from the concept of "readiness," which refers to a state of preparedness prior to entering the hostile environment.

As discussed in the body of the report, there has been virtually no systematic, empirical research into the "combat effectiveness" of units. To a large extent, this lack of research has been due to an inability to create an empirical measure of combat effectiveness -- a replicable, valid, coherent scale against which unit performance can be reliably gauged. Hence, the development of such a scale and the resulting improved understanding of the concept of combat effectiveness were the first goals of the research effort.

Given the ability to scale unit performance, it was then possible to examine associations between "effectiveness" and other variables -- tributes of units, levels of performance, enemy capabilities, and so forth. Using research techniques that blended the judgment of experienced officers with empirical evidence, the research team was able to create algorithms reflecting these associations. These algorithms, in turn, have implications for

- Procurement decisions,
- Allocation of training time,
- Research priorities,
- Unit composition, and
- Development of doctrine.

In addition, a particular effort was made to focus on the readiness implications of the research. Specifically, three concerns were examined:

1. To what extent can the research be used to find ways of projecting combat effectiveness before a unit is deployed in a potentially hostile environment?
2. To what extent can research be used to evaluate exercise performance an an indication of future combat effectiveness?
3. To what extent does the research validate existing systems for measuring readiness or suggest innovations in readiness measurement?

FOCUS

These objectives are ambitious. During the first year of effort, research focused on U.S. Marine Corps infantry battalions engaged in offensive operations. Marine Corps units were selected because

- The Commandant of the Marine Corps has a strong interest in the subject of combat readiness, as is evidenced by the recent development of the Marine Corps Combat Readiness Evaluation System (MCCRES) by Marine Corps personnel.

- Headquarters Marine Corps provided assistance in the form of an experienced officer to serve as the Contracting Officer's Technical Representative (COTR) for the study.

- Outstanding facilities of the U.S. Marine Corps Historical Branch and Library were available to support the data collection effort.

At the same time, every effort was made to ensure that the research was not focused exclusively on the Marine Corps, but on the general issues of ground combat. This is intended to facilitate application of the findings and, where appropriate, broadening of the research effort.

Infantry battalions were examined because they are the single most numerous units, and because battalions are the smallest units for which accurate data are likely to be found. The lowest level of analysis practical was desired in order to ensure focus on combat -- interaction with hostile forces. At the same time, data were collected on the activities of higher headquarters and supporting elements as their actions impacted on the infantry battalions being studied.

Focus was limited to offensive operations for two reasons:

1. Some substantive focus was necessary to produce a coherent data set for analysis within the limited resources available.

2. Offensive operations are among the most frequent and important undertaken by U.S. military units and are emphasized in current doctrine.

Specific military operations, usually confined to a few days, comprised the units of analysis. Twenty-two cases from four combat eras were examined -- World War II (10 cases), Korea (5 cases), Vietnam (5 cases), and special operations (2 cases; Dominican Republic and the 1958 Lebanon landings). Cases were chosen to provide a variety of experience.

APPROACH

The research approach was to tap the judgment of officers with combat experience. Rather than asking officers to list the factors which they believed to be most important in determining combat effectiveness, however, the study team wanted to create a "context." These contexts consisted of brief (8-21 pages) narrative descriptions of specific engagements. There were three reasons for using this format:

- Other researchers have in the past, and are today, using survey instruments to gather the abstract opinions of experienced officers on this topic.
- Collection of opinions in the abstract has a strong tendency to reproduce doctrine -- officer's opinions are shaped by training as well as experience, and the training is often more recent.
- Use of historical cases allows cross-checking of officer opinion with the historical information available.

The officer participants were asked, based on their readings of the narratives, to provide information on three issues:

1. Did the unit accomplish its mission?
2. Compared with other cases, how well did the unit accomplish its mission?
3. Why did the unit succeed or fail -- what were the crucial factors that contributed to the unit's relative level of mission accomplishment?

This "judgmental" data were supported by "historical" information coded by the project team. Historical data were collected for factual issues (how many personnel were in the unit, what were their levels of experience, and so forth) or those that required detailed examination of historical records (maintenance problems encountered, availability of specific types of ammunition, and so forth).

The project involved, then, three major data sets:

1. Judgmentally derived codings of mission accomplishment that reflected relative combat effectiveness.

2. Judgmentally derived data to determine which factors were crucial in determining whether units would perform well or poorly.

3. Historical data about the engagements.

The research results were derived from an analysis of these data sets and the associations among them.

RESEARCH RESULTS

Caveats -- Limitations on the Findings

No research effort can have universal application. Indeed, well-designed research is usually narrowly focused to reduce the number of confounding variables present, allow greater confidence that the findings are meaningful, and to increase understanding of the processes at work.

The major substantive limitations of the analyses presented below can be summarized in two main points:

1. The data are historical, not focused on the future. Hence, characteristics of the future battlefield are either underrepresented or not present at all in the data set.

2. The focus is on Marine Corps infantry battalions in offensive missions. Extension to other types of units, levels of command, services, or classes of mission can only be made with great caution.

From a statistical standpoint, the reader must understand four principal limitations on the interpretation of the analyses. First, the analyses are cross-sectional, not over time. This means that they are based on differences among the 22 engagements, or the 50 critical factors coding sheets, the 68 ratings of combat effectiveness, or the 101 comparative rankings of combat performance, not on repeated observations of a single unit. To be truly valid, the findings should be tested by following units over time.

Second, the analyses are limited by the historical record available. Neither the narrative descriptions nor the objective data collected by the study team are complete or perfectly accurate. The presumption is made that there are no consistent biases in the data set so that errors of omission are part of "random error."

Third, the absence of findings regarding a variable or class of variable indicates that no evidence was found indicating that the variable was systematically associated with changes (variance) in levels of combat performance. Hence, the analysis can omit factors or functions that are important in determining combat outcomes, but which have adequate values in all or nearly all engagements under study.

Finally, and perhaps most important, are the issues of numbers of cases and variance in the dependent variable. Statistically, 22 cases is a very small number with which to work. The results of the analyses are remarkably strong in light of this problem, but it did present difficulties during the analytic phase of the research. It was made even more serious by a lack of variance in the measures of combat effectiveness (see Chapter 4).

Having noted these problems, there are two further points to be made. The findings make excellent empirical and theoretical sense. Nevertheless, readers and potential users should be aware of these caveats and limitations. Second, based on the success of the methodology, a second research effort is planned that will directly resolve many of these problems. The next phase will include

- Analyses of unit performance from the 1973 war, where the "emerging threat" is better represented.
- Analyses of cases in which U.S. Marine Corps battalions encountered short effects and/or surprise from enemy force levels, positions, weapons, or tactics. This should
  - Provide more cases of imperfect performance, thus increasing the variance in the dependent variable.
  - Allow projection of the attributes of units that are prepared to withstand shock and surprise, a key element in the early days of conflicts.
- Analyses of units over time as well as cross sectionally.
- Enlargement of the data set and collection of variables that are focused to resolve ambiguities and uncertainties resulting from the first research effort.

These caveats apply, with different force, to all five sets of analyses reported below:

- The coherence and meaning of the judgmental measures of combat effectiveness.
- The coherence and meaning of the judgmental data regarding critical determinants of combat outcomes.
- The pattern of association between judgmental critical determinants and combat effectiveness.
- The pattern of association between historical data and combat effectiveness.
- The potential for innovations in the evaluation of exercises.

Measuring Combat Effectiveness

The first goal of this project was to develop a reliable, valid way of measuring combat effectiveness. Table 1 shows the results of a "factor analysis" performed on the 13 different judgmental measures coded. High (approaching 1.00) loadings and communalities are shown for all variables. These loadings, plus the failure of any other statistically strong factors to emerge, indicate that

- Judgmental measures of combat effectiveness form a coherent, unidimensional scale for the 22 engagements examined.

Of the 13 questions asked of the panel, the "marker" variable -- that variable that is best used as a summary indicator -- is a 10-point ranking which focuses the coding officer on mission accomplishment and ignores extenuating circumstances.

- Mission accomplishment is a central consideration in the minds of experienced officers asked to evaluate the relative combat effectiveness of units.

Several other findings emerged from a careful examination of the judgmental data regarding combat effectiveness:

- Units are rated less effective when overall "combat effectiveness" and "military mission accomplishment" are evaluated in terms of the unit in isolation. They are rated more effective when contextual factors and/or the performance of superior headquarters and/or support units is included in the analysis.

This can be read as reflecting common sense -- that an infantry battalion does not operate in a vacuum, but rather in a world of obstacles and constraints. More significantly, however, it also implies that adaptability is an important element in determining performance. The ability to react to different levels of resource availability, to take advantage of

TABLE 1

Results of Factor Analysis on Judgmental Combat Effectiveness

Question Number	Factor Loading	Estimated Communality
1	-.89	$\boxed{.91}$
2	$\boxed{.90}$	$\boxed{.92}$
3A	$\boxed{.90}$	.85
3B	-.81	.84
3C	$\boxed{.90}$	$\boxed{.91}$
4A	$\boxed{.90}$	.88
4B	(.76)	(.78)
5A	.84	.88
5B	.83	.89
6A	(.76)	.86
6B	.86	$\boxed{.92}$
7A	.88	.84
7B	.88	.80

Eigenvalue = 9.77

$\boxed{}$ = highest (over .90)

( ) = lowest (below .80)

opportunities provided by terrain, and to react to enemy resistance would be expected, based on this finding to be an important discriminator between successful and unsuccessful units.

- No bias was detected in coding for the different eras of combat -- the mean values for the four groups (World War II, Korea, Vietnam, and special operations) are remarkably close.

- "Objective" or historical measures of combat effectiveness are costly and time consuming to collect at the level of infantry battalions and do not vary widely across cases.

- Friendly casualties are not a good measure of combat effectiveness.

- Combat effectiveness is a multidimensional phenomenon that is not accurately reflected in any single objective number, but can be captured through integrated judgments.

Judgmental Evaluations of Critical Factors

A factor analysis was also performed on the set of 13 judgmentally derived critical factors variables available for analysis. Table 2 shows the results of that analysis. The three "factors" represent clusters of related variables. The association between each variable and each factor is measured by the factor loadings. A loading of 0.30 or greater shows reasonable association (1.00 is perfect). The communality refers to the association between the variable and the entire set of factors (the mathematical solution of the factor analysis problem).

Table 2 indicates that there are three statistically distinct dimensions in the data. Careful review of the pattern of variable loadings indicates that these can be labeled

- Coordination Functions (Factor I),
- Command and Planning Functions (Factor II), and
- Supporting Fires (Factor III).

TABLE 2

Factor Analysis of 13 Judgmental Critical Factors Variables

Factor Loadings

Variables	Factor I	Factor II	Factor III	Communality	n
Quality of Information	$\boxed{.89}$	.38	.06	$\boxed{.93}$	44
Quality of Plan	.54	$\boxed{.69}$	.11	$\boxed{.78}$	47
Logistics Support	$\boxed{.48}$	.33	.20	.38	36
Awareness of Enemy Capabilities	$\boxed{.74}$	.34	.20	$\boxed{.70}$	42
Implementation of Principles of War	.31	$\boxed{.75}$	.22	$\boxed{.71}$	36
Maneuver During Action	.33	$\boxed{.61}$	.24	.54	46
Artillery Support	$\boxed{.58}$	.15	.31	.45	34
Naval Gunfire	.26	$\boxed{.65}$	.54	$\boxed{.78}$	21
Preparatory Air Interdiction	.25	$\boxed{.85}$	.27	$\boxed{.87}$	16
Close Air Support	.18	.22	$\boxed{.96}$	$\boxed{.99}$	17
Armor Support	-.22	.31	$\boxed{.76}$	$\boxed{.72}$	42
Linkages to External Units or Commands	$\boxed{.48}$	.32	.38	.47	42
Communications	.47	.15	$\boxed{.70}$	$\boxed{.74}$	38
Suggested Interpretation	Coordination Functions	Command and Planning Functions	Supporting Fires		

$\boxed{}$ = Highest loading of the variable on any factor.

○ = Loading, other than highest, of 0.30 or greater.

Equally important, the existence of these three statistically distinguishable factors greatly facilitates later analysis and the interpretation of findings. Units that perform these activities well are expected to be effective in combat. Here, again, the importance of <u>adaptive behavior</u> -- coordination and command activities <u>during</u> the engagement -- is underscored.

Finally, the selection of critical factors by experienced officers produced an interesting insight into the breadth and depth of battlefield vision of Marine Corps combat-experienced officers.

- Experienced infantry officers identify a wide-ranging, theoretically balanced set of factors as explanations of success and failure in combat.

<u>Association Between Judgmentally Derived Critical Factors and Combat Effectiveness</u>

A first effort at exploring the pattern of association between judgmentally derived critical factors and combat effectiveness was a factor analysis, identical to the one reported in Table 2, but including the mission accomplishment based, 10-point scale of combat effectiveness. Table 3 shows the results. The three fundamental factors remain intact. Analysis focuses on the pattern of association (loadings) of the combat effectiveness variable across the factors.

- Combat effectiveness is associated quite strongly with the effective execution of command and planning functions, moderately strong with the availability and use of supporting fires, and somewhat with coordination functions.

More detailed analysis of the same information was undertaken by examining the bivariate correlations between each of the 13 judgmentally derived critical factors and the marker variable for combat effectiveness.

TABLE 3

Factor Analysis of Judgmental Critical Factors and Judgmental Combat Effectiveness

Factor Loadings

Short Title	Factor I Coordination Factors	Factor II Planning and Command Functions	Factor III Supporting Fires	Communality
Combat Effectiveness	-.13	-.76	.36	.73
Quality of Information	.88	.30	.09	.87
Quality of Plan	.66	.51	.16	.73
Logistics Support	.54	.23	.23	.40
Awareness of Enemy Capabilities	.81	.22	-.17	.74
Implementation of Principles of War	.31	.89	.18	.92
Maneuver During Action	.33	.71	.20	.65
Artillery Support	.58	.09	.33	.45
Naval Gunfire	.41	.45	.59	.72
Preparatory Air Interdiction	.41	-.69	.32	.75
Close Air Support	.19	.20	.94	.95
Armor Support	-.20	.30	.77	.73
Linkages to External Units or Commands	.46	.32	.37	.45
Communications	.40	.20	.68	.65

These coefficients are shown in Table 4. Negative correlations are the predicted direction. Perfect correlation would be -1.00. The absence of association would be 0.00. Levels of significance of 0.05 or less indicate statistical association. Hence, it is quite reasonable to conclude that

- Judgmental codings of critical factors associate strongly with judgmental codings of combat effectiveness.

This same information is reinterpreted in Table 5 to produce estimates of the relative importance of each of the judgmental critical factors in determining combat effectiveness.

The findings based on these analyses can be stated at two different levels. At the highest levels of abstraction, they suggest that "move, shoot, and communicate" is not a bad mandate for infantry units in the U.S. Marine Corps. Indeed, there would be grounds for real concern if, at these higher levels of abstraction, the research did not produce findings consistent with the previous experience of senior officers and the general theories of offensive action that have been developed in the past.

There are, however, some distinctive findings at even higher levels of abstraction:

- Supporting fires, when they are involved in an action, are likely to be extremely important, but they are associated strongly with only one element of the infantry battalion's own activities -- communications.
- Planning, command, and coordination are very tightly intertwined. The unit that cannot carry out these activities simultaneously will almost certainly fail.
- Planning and command during an engagement dominate the associations with combat effectiveness. Supporting fires and coordination are each perhaps one-quarter[1] as important as this set of functions.

[1] Based on the relative size of beta coefficients in the regression analyses carried out using marker variables from the three major factors.

TABLE 4

Bivariate Correlations Between Judgmental Critical Factors and Judgmental Measures of Combat Effectiveness

Variable	Correlation With Combat Effectiveness	Number of Cases	Level of Significance	Percent of Variance Explained
Quality of Information	-.39	44	.005	15
Quality of Plan	-.47	47	.001	22
Logistics Support	-.29	36	.041	9
Awareness of Enemy Capabilities	-.25	42	.055	6
Implementation of Principles of War	-.85	36	.001	73
Maneuver During Action	-.66	46	.001	44
Artillery Support	-.30	34	.042	9
Naval Gunfire	-.52	21	.008	27
Preparatory Air Interdiction	-.68	16	.002	46
Close Air Support	-.59	17	.006	35
Armor Support	-.42	31	.009	18
Linkages to External Units or Commands	-.38	42	.C07	14
Communications	-.51	38	.001	26

TABLE 5

Approximate Importance of Judgmental Critical Factors in Determining Combat Effectiveness

Variables Important in Most Engagements

Short Title	Associated Factor	Approximate Weight[a]
Implementation of the Principles of War	II	7
Maneuver During Action	II	4
Communications	III	3
Quality of Plan	I	2
Armor Support	III	2
Quality of Information	I	2

Crucial Variables That Are Coded Less Than One-Half the Time

Short Title	Associated Factor	Approximate Weight[a]
Preparatory Air Interdiction	II	5
Close Air Support	III	4
Naval Gunfire	III	3

Variables Coded as Important, But Not Strongly Associated With Combat Effectiveness

Short Title	Associated Factor	Approximate Weight[a]
Linkages to External Units or Commands	I	1
Logistics Support	I	1
Artillery Support	I	1
Awareness of Enemy Capabilities	I	0.5

---

[a] Based on bivariate (Pearson product moment) correlations with judgmental combat effectiveness in comparison with 10 typical engagements. 1-10: Percent of variance explained.

- Coordination functions are an important element in determining combat effectiveness, but neither theoretically nor statistically do they become significant unless there is high-quality planning and command.

Dropping down one level of abstraction, it is also possible to draw conclusions at the level of the 13 individual variables drawn from the critical factors data set. Table 5 shows a breakdown of these variables and approximate weightings or relative priorities among them based on the percentage of variance that each variable explains in combat effectiveness. Many of these findings are important.

- The single most important variable, "implementation of the principles of war," is a composite referring to actions that the unit executes on the battlefield after an engagement begins. Being prepared to execute and react is vital. Here again, adaptive behavior appears to be central.

- The single most important function for unit success is maneuver during the action.

- Nonorganic supporting fires -- preparatory air, naval gunfire, and close air support -- are absolutely vital when they are involved in an action. Units must be trained to use them effectively if they are to achieve combat effectiveness.

- Communications are the second most important specific functions that an infantry battalion must perform well to operate effectively in combat. Units that communicate well also have a good record in use of supporting fires, although specific linkages to external units or commands do not show up as critical in themselves.

- Quality of planning and quality of information are important contributions to combat effectiveness. They are perhaps four times as important as awareness of enemy capabilities. This probably means (a) that there are not a large number of cases in the data set in which awareness of enemy capabilities was poor, and (b) that the information and planning functions depend on knowledge of the entire situation -- terrain, weather, enemy, disposition of own forces, and so forth -- rather than on knowledge of the enemy situation.

- Effective use of armor support is an important contributor in slightly over one-half of the cases analyzed. Emphasis on armor support in training would be an important element in preparing infantry battalions for combat.

- There is evidence that logistics support and artillery support have a positive impact on combat effectiveness, but there is little evidence that they have been frequent determinants of combat outcomes.

Association Between Historical Variables and Combat Effectiveness

Analyses of the patterns of association between historical variables and combat effectiveness were constrained severely by the small number of cases under analysis. This problem was particularly difficult in these analyses because there were no appropriate techniques available for artificially increasing the size of the data set. Each case could yield only one observation, in contrast with the critical factors data where multiple codings of the same case were used in the analysis. In addition, the small number of cases of poor performance made it difficult to generate sufficient variance to perform standard associational analyses.

To deal with these difficulties, the research team

- Created variance by working with data selected to reflect "above average" and "below average" combat performance, rather than mission accomplishment, and

- Cross-checked the analyses of the whole set of cases against a subset of nine "extreme cases" which were found to be universally ranked above average or below average.

Conclusions Based on Historical Data

The historical data viewed in isolation provide fewer insights than the critical factors data. There are, however, some new findings. It is important to remember that these findings are based on a comparative standard of combat effectiveness, rather than the absolute one used with the critical factors data. That is, the best cases of performance are being compared with the worst, while "adequate" performance is not measured by the scale.

- Completion of a full cycle of unit training before commitment to a combat environment increases the probability of effective performance.

- Air support during an engagement increases the probability of effective combat performance of infantry offensive missions.

- Loss of internal contact among the components of an infantry battalion decreases the probability of satisfactory combat performance.

- U.S. Marine Corps infantry battalions have, in the cases studied here, performed less effectively when facing intense enemy artillery, intense enemy mortar fire, and enemy armor than when those factors were absent.

- Intense enemy infantry resistance has led to increased combat effectiveness by U.S. Marine Corps infantry battalions, perhaps by fixing the enemy in position.

- Marine Corps infantry battalions have performed more effectively against mixed and irregular enemy forces than against veterans and regulars.

- With the element of surprise on their side, U.S. Marine Corps battalions have had an increased probability of success, while they have generally been able to neutralize enemy tactical surprise situations.

- U.S. Marine Corps battalions composed of a mixture of regulars and veterans with replacements and reservists have performed as well as or better than regular and veteran units without reservists or replacements.

- Neither regimental-level training, division-level training, nor rehearsals for the specific engagements show a positive association with effective combat performance. The data suggest that they may detract from probability of satisfactory performance, perhaps by distracting the unit from more fundamental training.

- Prepared enemy positions have not caused lower probabilities of infantry performance. Like the intensity of infantry fighting, they may fix the enemy in a position where it can be destroyed by fire and maneuver.

- Artillery support during an engagement does not increase the probability of successful performance by a unit. Support is most intense during difficult combat. There appears to be room for improvement in doctrine and employment of artillery.

- Artillery preparations are negatively associated with effective performance. This may reflect the fact that difficult offensive missions are often preceded by heavy preparatory fires. However, these fires are not effective enough to bring up the probability of success to an equal level with other engagements.

- Moderate and light air preparations are likely to increase the probability of satisfactory effectiveness over those of engagements where no preparatory air is involved. However, heavy air preparations, associated by definition with difficult combat, do not in themselves increase the chances of successful combat.

- Ammunition expenditure, artillery ammunition availability, and mortar ammunition availability are not found to be either important aids or hindrances to effective combat performance, although consumption is higher during intense combat.

- Supply and delivery of supplies are not found to be either a major problem or a major determinant of combat effectiveness.

Exercise Evaluation

The research team also compared its findings with information systems such as the Marine Corps Tactical Warfare Simulation Evaluation and Analysis System (TWSEAS) and the Marine Corps Combat Readiness Evaluation System (MCCRES), and exercise evaluation documents from a variety of Marine Corps commands at all levels. The research issue of interest was potential for innovations based on the combat effectiveness research effort.

There is an opportunity imbedded in this system for developing a different way of evaluating units. The following discussion is predicated on a desire to use that opportunity. Basically, the argument is as follows:

- Exercises are opportunities for both learning and evaluating.
- The research reported in this volume indicates that the "adaptive behavior" of a U.S. Marine Corps infantry battalion is the most important component of combat effectiveness.
- Learning is a form of adaptive behavior.
- CACI researchers found that is is possible to project combat effectiveness based on a relatively small number of unit functions.
- Therefore, it is both fair and wise to collect data relevant to absolute levels of performance on those indicators <u>and</u> on the learning (adaptive) behavior of the unit over the course of the exercise.

In other words, it is possible to evaluate unit potential for effective combat performance both on the basis of absolute performance and on improvement rates observed over time. This approach will not completely eliminate the problem of "report cards." It would, however,

- Force the use of more objective indicators, making identification of marginal areas of performance easier,
- Provide an opportunity to evaluate the unit for adaptability, a crucial area currently not examined, and
- Reduce the incentive for underreporting exercises both by creating more objective data and allowing demonstration of the learning curve for units during the exercise.

A visit by the principal investigator to the Marine Corps Air-Ground Combat Training Center (MCAGCTC) at Twentynine Palms, California, and observation of field exercises on that base suggested that a key element was being added to the exercise system in the Marine Corps -- the idea of comparability. A series of exercises involving live fire (and therefore no major live aggressor play) has been conducted at Twentynine Palms for battalions from both the East and West Coasts. All battalions have been given roughly the same mission, all faced the same terrain, and, with some variation, all have encountered the same scenario obstacles.

This element of comparability, particularly in light of the existence of TWSEAS technology, suggests that the Marine Corps is in a unique position to establish an evaluation system for projecting combat effectiveness from exercise data. Such a system would have two principal modes of implementation:

- To produce a baseline data set, validate the concept, and gain experience in its implementation, a phase in which a single, replicable exercise is evaluated, according to experimental design and quasi-experimental design principles, and
- A later phase in which the assumptions of the design are relaxed and data from different exercises, missions, terrain, and other key features are collected and compared with the baseline system.

Since units will experience the Twentynine Palms problem over time, however, new and different elements are constantly entering the problem. For example, new threats emerge, new techniques or weapons may be available to the exercise battalion, levels of support may be different, and the weather will certainly be different for each unit. Rather than classical experimental design (ideal for running rats through mazes), reliance on quasi-experimental design (Campbell and Stanley, 1963) techniques is wise. These techniques call for creating a baseline of data on key variables, and then performing both simple statistical associations and more elaborate multivariate techniques to establish "causal relationships."

Any data from any source can be utilized. Scores from the MCCRES system would make an excellent starting point since MCCRES is both comprehensive and standardized throughout the Marine Corps. However, there is a great deal of highly structured information currently passing through the hands of umpires and the Troop Exercise Control Center (TECC) that could be extremely valuable. For example, interviews at Twentynine Palms indicate that umpires are currently recording, for each target attacked,

- The type and time of intelligence information received, which indicates the presence of a target,
- The time at which the target is acquired by the unit,
- The time at which fire is brought on the target (direct or indirect), and
- The time at which the target is declared neutralized, that is, hit by sufficient firepower to destroy it.

Figure 1 shows the type of data that might be generated from this information. The horizontal axis reflects hours into the exercise. The vertical axis is the time from the identification of a target by the unit (target acquisition) until the target is declared destroyed by the TECC. Note that these data reflect two things, the relative skill level of the unit at the time the exercise began (needing 8-10 minutes to destroy a target), and the learning or adaptive behavior of the unit. Given the importance of factors like maneuver during action, implementation of the principles of war, use of supporting fires, and so forth, in the research reported earlier, this type of learning behavior is an outstanding indicator of the quality of unit performance.

It is not the purpose here to design a full evaluation system, but a brief list of the types of key indicators that are directly available and could be coded for units would include

- Speed of mission completion (where missions are comparable),
- Speed of response to enemy air threat,

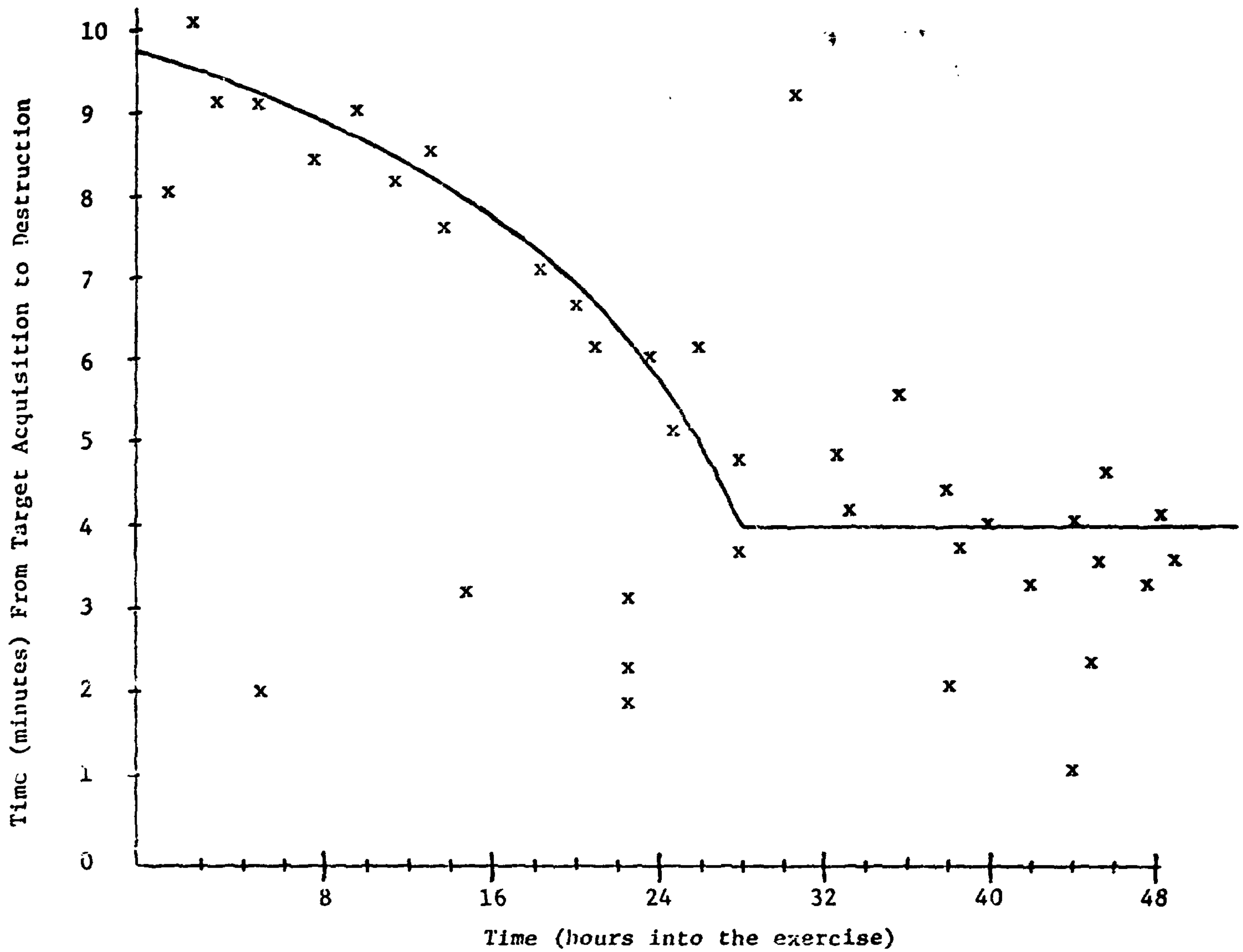

Figure 1. Illustrative Plot of Performance Data

- Variety of weapons utilized,
- Reaction (percentage success and speed) to electronic warfare,
- Speed of destruction of enemy armor, and
- Percentage of time out of communication with supporting arms.

Two key types of related analyses would also become possible. First, weighted performance scores could be produced. Second, and perhaps more important, research into the linkages between exercise performance measures and other variables would be practical. For example,

- The finding that completion of full unit training cycles is a predictor of combat performance could be examined for exercise situations;
- Unit maintenance ratings could be compared with exercise performance measures;
- Personnel turbulence variables could be used to estimate the impact of rotational policies on performance levels; and
- Number of weather-related injuries might be correlated with unit performance to establish overall levels of preparedness.

These analyses necessarily depend on the gradual development of a large enough data base to perform meaningful statistical analyses. Once sufficient information is present for these baseline data, comparisons with other types of exercises, perhaps already collected through TWSEAS or another system, could be used to produce evaluations. This "system" is, then, a collection of techniques and approaches already in existence but structured for effective analyses and weighted on the basis of research into effective combat performance.

THE STRUCTURE OF THE REPORT

Like all serious research efforts, this project was complex. Its findings are based on dozens of research decisions and assumptions. The main body of the report should serve to make these explicit and provide supporting data and arguments for the brief summary presented here. Chapter 2 is a detailed introduction to the research and focuses on its rationale and research design. The data collection procedures utilized are reviewed in Chapter 3. Chapter 4 covers analyses of the combat effectiveness measures in isolation. The judgmentally derived critical factors data are analyzed in detail in Chapter 5. The association of both the critical factors data and the historical data with combat effectiveness measures is reviewed in Chapter 6. Finally, Chapter 7 covers the suggested innovations in exercise evaluation procedures.

The reader who is interested primarily in projecting combat effective performance may want to confine his/her initial reading to Chapters 6 and 7. Insight into the concept of combat effectiveness and the patterns among the judgmentally derived critical functions can be generated by reading Chapters 4 and 5.

CHAPTER 2. INTRODUCTION

BACKGROUND

One of the most crucial problems facing military commanders and defense decision-makers is the maintenance of an "effective" military capability. Creating combat-ready units in time of war, establishing standards and priorities for training, procuring, and staffing during periods of relative peace, and assuring a military capability that deters potential adversaries from dangerous adventures have long been central missions of the U.S. armed forces. Yet, measuring "military effectiveness" without data from combat experience remains a major unresolved research problem.

Military operations are complex undertakings involving a variety of social processes (such as C^3, planning, and so forth), physical movements, mechanical operations, and rapidly changing situations about which knowledge is always incomplete. These have traditionally been described in terms of functional areas -- command, operations, logistics, intelligence, and so forth. The failure of any one functional area may undermine the performance of a unit ("for want of a nail a shoe was lost..."). But military effectiveness is more than the sum of its functional parts. All functional aspects of an operation may go well, but a unit may still fail to accomplish its mission. Conversely, history is replete with examples of military units that achieved their missions despite functional failures.

The U.S. military has long recognized this and has therefore made commanders responsible not for mere technical proficiency within their units but for effective performance. Units that have fewer functional failures, overcome obstacles, and achieve their objectives are viewed as "successful." Focus is placed on outcomes -- the achievement of desired goals -- rather than the process involved in getting there. In short, military effectiveness can only be measured in terms of mission accomplishment.

The system currently used to estimate the status of military forces focuses on unit readiness. Readiness is indicated by a number of unit "attributes" that are measured either objectively or subjectively. Hence, completion of required training hours, availability of crucial items of equipment, condition of equipment, estimates of unit morale, and a myriad of other tangible and intangible variables are considered in estimating a unit's level of readiness. Unfortunately, this system has a number of disadvantages. For example,

- Some of the measures are sterile and viewed as meaningless by many commanders.
- The intangible variables are estimated subjectively, and wide disagreement can occur among qualified commanders.
- The system lacks historical validation.
- Reliance on specific "attributes" creates incentive for unit commanders to emphasize particular areas and downplay others.
- The relationship among the components of "readiness" is unclear, so a commander is given relatively little guidance as to priorities in upgrading unit capabilities.
- The system is relatively insensitive to improvement -- when a unit is deemed "ready," no upgrading is possible, regardless of new equipment or training.

"Readiness" refers to the level of preparedness that a military unit achieves before being deployed in combat or an area where combat may begin at any moment. "Effectiveness," as used in this report, refers to a unit's level of performance after it has been deployed in combat or in an area where combat is possible. Ultimately, this research effort is intended to contribute to the development and validation of readiness measures that will project effective combat performance. However, the first-year research reported here does not include analyses of the current readiness system. Rather, this research begins with the concept of "mission accomplishment" and seeks to employ a combination of experienced military judgment with appropriate quantitative techniques to create a reliable, valid system for measuring "military effectiveness."

The research is supported by the Defense Advanced Research Projects Agency (ARPA) with the cooperation of the U.S. Marine Corps. It focuses on a single type of unit, the Marine Corps infantry battalion, but is designed to facilitate expansion to include other types of Marine Corps units and/or other services.

OBJECTIVES

One major objective of this first year of effort is to demonstrate the soundness of the overall concept of "mission accomplishment" as an approach to measuring combat effectiveness. In addition, the research is intended to produce

- A measurement tool for evaluating the military effectiveness of units in future exercises and operations,
- A set of priorities and weights reflecting the relationships among unit attributes, specific military functions, and mission accomplishment,
- An estimate of the linkages between relatively "quantifiable" performance measures and judgmental measures of military effectiveness, and
- An explicit set of assumptions and criteria by which experienced officers evaluate mission accomplishment and measure their agreement when these values are utilized.

These products have a wide range of specific applications in the Marine Corps and potentially in other services. The establishment of priorities in key areas such as training, doctrine, procurement, and research would be greatly enhanced. Moreover, the creation of a replicable, validated set of measures for military effectiveness would produce a powerful management tool and a sound basis for making and defending budgetary decisions and requests. For example, the newly developed Marine Corps Combat Readiness Evaluation System (MCCRES) has been formulated on the basis of experienced officer judgment. Findings from the empirical research reported here can be used as a validation for some parts of that system and may well cause reanalysis and rethinking of other elements within it.

In addition, some of the material produced for the report may be useful in itself. For example, the Marine Corps is considering using the 22 narrative descriptions of military operations in its officer school system. Similarly, the list of functional activities occurring in combat, developed for data coding during the project, is a contribution to the literature in itself.

METHODOLOGY

This project is being carried out within a larger ARPA research program. Other research in the program is focused on using decision analysis and hierarchical inference structures to identify functions and attributes that Marine Corps officers believe cause units to be effective during combat and to prioritize those areas. The CACI, Inc.-Federal approach is to attempt to tap that same experience, but to do so by creating a context for the judgment, a data base about the context, and comparisons between the judgmental codings and traditional measures of combat effectiveness such as casualties taken and given, territory won or lost, and resources expended.

The basic research strategy, depicted in Figure 1, was to employ a "quasi-experimental" design in which two types of predictor (or independent) variables were used to create an algorithm that explains a single dependent variable (mission accomplishment). Mission accomplishment was estimated by expert judgment using techniques tested and proven by CACI on past ARPA projects, including modified Delphi approaches. The best of the available quantitative measures of predictor variables, including unit attributes, measures of performance, and measures of residual capacity, were used, along with a set of expert judgments about the reasons for mission accomplishment or failure, to develop a balanced system of predictor (or independent) variables. Controls regarding the type of situation and mission were also used. The resulting predictor system was then subjected in a variety of validation exercises, both quantitative and qualitative, before a system for estimating military effectiveness was developed.

The power of this approach lies in its simple, direct confrontation with the issue of military preparedness. Factors that are shown to be associated with mission accomplishment and mission failure are identified as the critical aspects of performance. These, then, become important items in "explaining" or "predicting" unit performance.

STRUCTURE OF THE REPORT

Chapter 1 comprised the Executive Summary of the report. The next chapter (Chapter 3) reviews the data collection effort, including cases selected, judgmental information, and historical research. Chapter 4 reviews the data set on unit performance, mission accomplishment, or "combat effectiveness." Judgmental data regarding factors that appeared crucial in the engagements are covered in Chapter 5. Chapter 6 covers associations between the independent variables (both judgmentally and historically derived) and the combat effectiveness scores. Suggested algorithms for "predicting" effectiveness are discussed in this chapter. A recommended approach to evaluating Marine Corps infantry battalions when participating in exercises is suggested in Chapter 7.

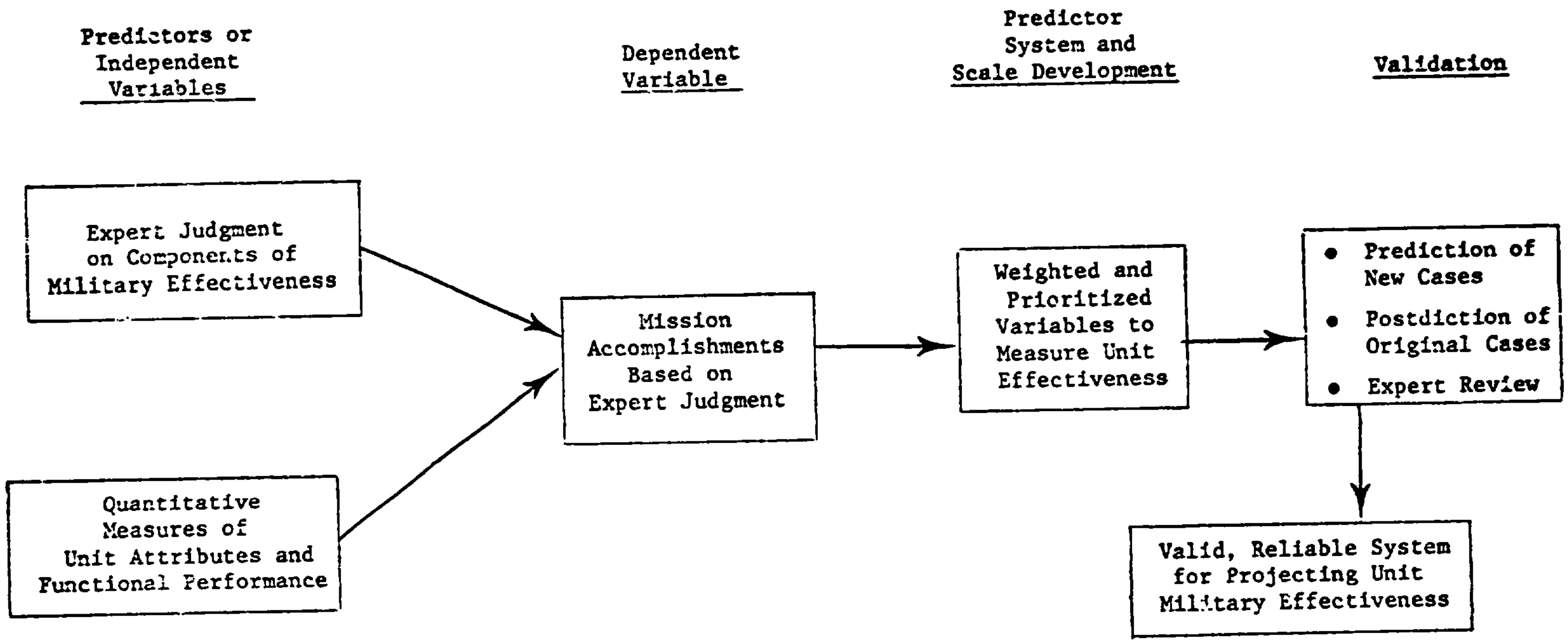

Figure 1. Basic Research Strategy

The voluminous material produced to execute the contract has been placed in Volume II. Appendix A of that volume contains the 22 narrative descriptions of operations utilized in the study. Appendix B is the questionnaire used to elicit combat effectiveness ratings from officers who read the descriptions of operations. Appendix C is the questionnaire utilized to elicit judgment about the factors and functions that these officers felt were crucial in determining unit effectiveness. Finally, Appendix D includes the forms used by the CACI, Inc.-Federal research team to assemble historical data for analysis.

CHAPTER 3. DATA COLLECTION

OVERVIEW

This chapter reviews the four data collection efforts carried out during the research project. The first of these was case selection and development of narrative descriptions for selected military operations. The second and third efforts, accomplished simultaneously, involved collection of judgmental data about the case studies. These included "combat effectiveness" and the relative levels of unit performance in a variety of different combat functions. The last data set compiled by the CACI, Inc.-Federal research team includes structured materials on a variety of factual or "objective" unit attributes, the situation, enemy forces present, and the results of the engagement. Each data collection effort is treated below. The descriptions include procedures utilized, problems encountered, and examples of the criteria or coding forms utilized. The complete set of coding forms may be found in Volume II of this report.

NARRATIVE DESCRIPTIONS OF MILITARY OPERATIONS

Case Selection

Two terms of reference for the research provided some criteria for case selection -- focus on Marine Corps infantry battalions and availability of sufficient data to support the analyses. With the help of the U.S. Marine Corps Museum and Library staff, three further criteria were developed:

1. To facilitate identification of relevant documents and evaluation [illegible]cs operating independently or largely independently were favored.

2. The analyses were confined to units engaged in offensive missions. This provided a continuity to facilitate comparative analysis and was judged necessary in light of the relatively small number of cases to be utilized in this first effort.

3. Oversampling from Vietnam was avoided because the in-depth historical research on that era is still underway (denying the research team the advantage of validated secondary sources to help structure the data), because of the artificial nature of some reporting statistics, and because of the relatively unique battlefield conditions during that conflict.

Discussions with the Contracting Officer's Technical Representative (COTR) from Marine Corps Headquarters during the early weeks of the project produced additional guidance. The COTR indicated that the research effort would be most useful if it could include engagements that had key characteristics of the modern battlefield -- enemy antiair capability, enemy mechanized forces, and urban conflict. These are, of course, difficult attributes to match in historical combat for U.S. Marine Corps battalions, although some success was achieved.

Procedurally, the engagements were selected by reviewing the best available secondary sources to locate potential cases, by discussions with personnel at the Marine Corps Museum and Library regarding the extent of primary material available and secondary research already carried out on the engagements, and by a review of documents available in the Washington, D.C., area. Data availability, as might be expected, proved to be the single most important factor.

Table 1 shows the engagements included in the set and gives a brief summary of the major reasons for their selection. A mixture of missions between assault landings and other types of offensive operations was consciously sought. DEWEY CANYON and the two Khe Sanh engagements include significant enemy antiair capabilities. Yongdungpo involves enemy mechanized forces and urban fighting. Hue City, Seoul, and the Dominican

TABLE 1

Cases Utilized

ID Number	Group	Unit	Short Title	Unit Mission	Reason for Selection
1	WWII	3/1	Peleliu I	Amphibious Assault	Difficult Mission
2	RVN	2/4	STARLITE	Assault, Search	Heavy Contact
3	RVN	3/9	DEWEY CANYON	Planned Attack	Enemy Antiair
4	WWII	2/1	Peleliu II	Attack	Intense Combat
5	RVN	2/5	Hue City	Counter-attack	Urban Combat
6	Korea	3/5	Inchon	Amphibious Assault	Hasty Planning
7	RVN	2/3	Khe Sanh I	Search, Assault	NVA Opposition
8	RVN	3/3	Khe Sanh II	Search, Assault	Intense Combat
9	WWII	2/28	Iwo Jima, Suribachi	Amphibious Attack	Difficult Mission
10	WWII	3/25	Iwo Jima, North	Amphibious Attack	Intense Combat
11	Korea	2/1	Yongdungpo	Attack	Urban Anti-mechanized
12	Special	3/6	Dominican Republic	Land, Secure Area	Urban Mission
13	Special	2/2	BLUEBAT (Lebanon)	Land, Secure Area	Exercise Comparison
14	WWII	3/29	Okinawa, Motobu	Amphibious Assault	Difficult Mission
15	WWII	3/29	Okinawa, Oroku	Attack	Comparison with Case 14
16	WWII	1/3	Guam I	Amphibious Assault	Difficult Mission
17	WWII	2/9	Guam II	Amphibious Assault	Comparison with Case 16
18	Korea	2/1	Seoul	Attack	Urban Combat
19	WWII	1/29	Saipan	Attack	Difficult Mission
20	Korea	3/5	Yudam-ni Breakout	Attack	Difficult Mission
21	WWII	2/1	Cape Gloucester	Blocking	Independent Mission
22	Korea	1/7	JAMESTOWN	Attack	Reported Poor Performance

Republic are primarily located in urban environments. Variety was also sought in the time available for planning, the quality of enemy forces, the training and experience of U.S. Marine Corps battalions, and the types and amount of support requested and available.

Considerable attention was paid to the frequency of sampling from each category of engagement. World War II cases were generally the best researched and archived. Vietnam cases suffered from lack of previous research, atypical conditions, and some artificial reporting structures. Surprisingly, the Korean war material proved largely underresearched by historical scholars, and the archives for that conflict were often less complete than those for World War II. A direct trade-off was present, therefore, between the recency of available data and the probable quality.

A special problem was encountered in terms of exercise operations. The original research design called for including exercises in the data set. The idea was to compare results from combat engagements with peaceful operations and exercises that occur before units are actually deployed. Rather sustained efforts by the study team, with the cooperation of the Marine Corps Museum and Library, failed to identify any exercises documented in sufficient detail to support analysis. Two special operations, Dominican Republic and Lebanon, were included in the set to offset partially the lack of exercise data. In addition, these two operations represent a different type of mission, involving heavy constraints on actions, high tension, political implications, and potential for immediate fighting. It is one for which Marine Corps units have had to train and plan for in addition to main combat roles.

<u>Narrative Development</u>

Preliminary research to identify the amount and quality of information available was extended when a case was actually selected. The key issues at this point became bounding engagements in space and time, locating as many primary sources as possible through the archive system, identifying relevant secondary sources, and assembling all materials in one place. All sources, including oral histories and meetings with historians, were utilized. Where necessary, arrangements were made for appropriate declassification.

The CACI research team was led, in this phase, by two retired Army officers, one a Major General and the other a Colonel. Support was provided by a variety of other senior research associates (Ph.D.-level personnel) and research assistants. The narratives themselves range in length from 8 to 22 pages, and most include sketch maps of the action. Vietnam engagements were written with appropriate map sheets available to supply more relevant terrain information. The entire set of narrative descriptions is located in Volume II, Appendix A.

DATA COLLECTION PROCEDURES

Combat Effectiveness Judgmental Data

A crucial element in the research was the identification of unit performance (relative combat effectiveness) measures. These are not necessarily measures of readiness (levels of preparation), though some measures of readiness may prove to be indicators of combat effectiveness.

It was a basic premise of this study that unit effectiveness is best measured by tapping the judgment of experienced officers on the issue of "mission accomplishment." Regardless of the problems involved, the military functions involved, or the obstacles encountered, units that accomplish their missions are "effective" and those that do not are "ineffective." Military missions are relatively complex and include such components as delaying an enemy advance, seizing specific territory, breaking through a line of resistance, maintaining capability to support other units, and so forth.

The research design called for dealing with this complexity by tapping the experience and judgment of senior officers. Active duty and retired Marine Corps officers in a variety of locations assisted. Seventeen individual officers participated in data coding. They represented combat experience ranging from World War II through Vietnam. Some were volunteers from Headquarters, Marine Corps, others were awaiting assignment to service schools, some were volunteers from the retired ranks, and one group was arriving at the Naval War College for a year of schooling. All except one (an artillery officer) were infantry combat veterans. Over one-half had commanded battalions in combat. Standard biographical data were collected on the officer participants to allow statistical checks for bias.

Each officer was asked to read through a set of narrative descriptions. Upon completing each narrative, the officer filled out a questionnaire entitled "Overall Evaluation of Combat Effectiveness" (see Volume II, Appendix B for a copy of the questionnaire). Questions on this form were designed to elicit the officer's judgment of the relative mission accomplishment of the U.S. Marine Corps battalion described in the narrative. Question 2, which later proved statistically and theoretically most useful, read

> 2. Compare the overall combat effectiveness of the friendly battalion with that of units you have seen in combat. Using a 10-point scale, with 10 for the worst performance in the set and 1 for the best performance you have ever personally seen, rank this unit.

Other scales were tried, one requiring discrimination between 0 points for inadequate performance and 100 for outstanding, another asking for percentage of mission accomplished, and several requesting assignment of grades from A (excellent) to F (failure). These scales were used to ensure that as much variance as possible would be generated to support analyses, while at the same time avoiding data analysis that was unstable because of false precision.

Confounding circumstances were also introduced, including terrain, enemy resistance, support received, and the activities of higher headquarters. These factors were included so that statistical checks on the consistency of judgment could be made and sources of uncertainty about rankings identified.

In addition, officers were asked to rank the cases that they read to identify which unit performed better. Because the officers provided varying amounts of time to this task, different numbers of pair-wise comparisons were made. All in all, 101 such pairs were evaluated. Table 2 compares the frequency of opportunities for pair comparison with the frequency actually generated across groups of cases. Section A of the table shows the distribution of possible pairwise comparisons. For example, 19 percent of all possible comparisons would be between two different cases from World War II. Section B shows the actual number of comparisons made. A total of 12 comparisons were made between different World War II cases out of 101 total comparisons.

The distortions are deliberate. First, given the importance of the special operations (representing noncombat operations, high likelihood of

TABLE 2

Frequency of Paired Comparisons

A. Frequency of Opportunities for Comparisons (Percent)

	W	K	V	S
W	19	22	22	9
K		4	11	4
V			4	4
S				1

B. Actual Comparisons (n=101)

	W	K	V	S
W	12	22	15	8
K		12	12	4
V			12	3
S				1

Key: W = World War II V = Vietnam Conflict
K = Korean Conflict S = Special Operations

occurrence, and exercise data), virtually the full sample was taken there. Second, sufficient comparisons were made within the Korea-Korea and Vietnam-Vietnam pairings to ensure that it would be possible to discriminate between engagements during these conflicts. Third, given the relatively high-quality information available for the World War II periods, but its lack of currency, these data were used primarily as a basepoint for comparisons with other periods. Fourth, Vietnam data were used somewhat less because of atypical conditions and some artificial reporting structures.

At the completion of a coding period, which included the coding of critical factors discussed below, the officers in each group (ranging from two to three) were encouraged to compare the engagements they had read and evaluate them in light of previous rankings. In only one case did they disagree about these comparative rankings.

Critical Factors Judgmental Data

After completing the combat effectiveness evaluation, officer participants were asked to complete a form entitled "Judgmental Evaluation of Critical Factors" (see Volume II, Appendix C for a complete copy of this form). A typical page from this form is shown in Figure 1.

The instructions to officer participants were to fill in the form for the engagement under examination. The major headings (ALL CAPS) listed 39 different functions or activities drawn from the literature on combat performance. The subheadings (Initial Caps) were used to define the components to be considered in evaluating unit performance. The officer could check "no information" if, as was at times the case, the research team had been unable to provide sufficient material to allow clear judgment or if the category was not relevant to the engagement. If sufficient information was available, the officer was asked to indicate both the importance of the activity in the engagement and the rating he would assign. A sharp distinction was deliberately imposed between satisfactory and unsatisfactory performance.

	No Information	Importance to the Outcome of This Engagement	Rating: Unsatisfactory	Rating: Satisfactory
III.2.3 USE OF FIRE		low — high	low — marginal	marginal — high
a. Organic				
1. Fire discipline				
2. Location of organic crew-served weapons				
b. Timeliness of Organic Fire				
c. Accuracy of Organic Fire				
d. Fire Planning For Organic Weapons				
III.2.4 SUPPORT ARTILLERY				
a. Location of Trained Observers				
b. Communication With Supporting Units				
c. Fire Planning				
d. Use of Fire to Attain Unit Objectives				
e. Ammunition Availability				
f. Use of Appropriate Ammunition Type				

Figure 1. Example Page From Critical Factors Form

Table 3 lists the functions that officers were asked to code. Where groups are shown (three items under "Background to Engagement," for example) only the subsets were coded. In the two "Principles of War" questions it was possible to code the nine principles taken from Fleet Marine Field Manual 6-3, "Marine Infantry Battalion," individually. Table 4 shows the number of coding forms completed for each engagement.

After completing the day's coding, officer participants were asked to discuss the critical factors present in each engagement that they had coded, to suggest additional factors not present on the current form, and to compare their evaluations. These discussions were often lively, but fewer than 5 percent of them resulted in a desire to change an evaluation or add a new element to the list of critical factors.

Historical or "Objective" Data

There are a number of quantifiable features about any military operation that are not susceptible to a great deal of judgment -- the number of personnel present, experience, distances traveled, amounts of resources consumed, and so forth. In addition, other somewhat subjective, factual questions are impossible to convey in a relatively brief narrative description -- fire effectiveness, maintenance problems, and so forth.

The CACI research team was charged with researching these two types of information for each engagement. The complete set of data collection forms is assembled in Volume II, Appendix D. An example form, "Ammunition Variables," has been included as Figure 2. Research on these variables was begun during the development of narrative descriptions but extended well past their completion. Careful searches of archives ensured the best possible data coding. Titles of the 12 forms are shown in Table 5.

TABLE 3

Critical Factors

Background to Engagement

- Training
- Morale at Outset
- General Quality of Information

Overall Planning for Engagement

- Quality of Plan
- Consistency of Plan With Principles of War
- Timeliness of Orders
- Reserves
- Logistics Support
- Subordinate Units
- Awareness of Enemy Capabilities

Weather and Enemy Situation

- Weather
- Quality (Type and Experience) of Enemy Forces

Overall Effectiveness of Execution

- Implementation of Principles of War in Action
- Maneuver
- Use of Fire
- Support Artillery
- Naval Gunfire

Tactical Air Support

- Preparatory Interdiction
- Close Air Support

Overnight Positions

- Adequacy of Position
- Adequacy of Protection

Security on the Move

- Armor Support
- Linkages to External Units or Commanders
- Reaction to Unexpected Situations

Miscellaneous Factors

- Morale During Combat
- Discipline
- Aggressiveness
- Initiative
- Resourcefulness Under Pressure
- Casualty Levels (Impact on Effectiveness)
- Medical and Evacuation Support

Supply

- Ammunition Adequacy, Timeliness of Resupply, Availability of Key Types
- POL
- Food and Water
- Barrier Materials
- Special Equipment
- Maintenance -- Evidence of Failure in Action
- Communications

TABLE 4

Number of Coding Forms Completed
for Each Engagement

ID Number	Group	Unit	Short Title	Number of Critical Factors Forms Completed
1	WWII	3/1	Peleliu I	2
2	RVN	2/4	STARLITE	2
3	RVN	3/9	DEWEY CANYON	2
4	WWII	2/1	Peleliu II	3
5	RVN	2/5	Hue City	2
6	Korea	3/5	Inchon	2
7	RVN	2/3	Khe Sanh I	1
8	RVN	3/3	Khe Sanh II	3
9	WWII	2/28	Iwo Jima, Suribachi	2
10	WWII	3/25	Iwo Jima, North	2
11	Korea	2/1	Yongdungpo	2
12	Special	3/6	Dominican Republic	2
13	Special	2/2	BLUEBAT (Lebanon)	3
14	WWII	3/29	Okinawa, Motobu	2
15	WWII	3/29	Okinawa, Oroku	2
16	WWII	1/3	Guam I	3
17	WWII	2/9	Guam II	3
18	Korea	2/1	Seoul	2
19	WWII	1/29	Saipan	2
20	Korea	3/5	Yudam-ni Breakout	2
21	WWII	2/1	Cape Gloucester	3
22	Korea	1/7	JAMESTOWN	3

ENGAGEMENT NO. _____

Operation (Name, location, year) ____________________________________

Phase ______________ Dates ___________________ to ________________

Unit (Bn, Regt, Div) ____________________________________

AMMUNITION VARIABLES

1. Ammunition Availability:

	Ample	Sufficient	Marginal
Small arms (rifle, pistol, grenades)	_____	_____	_____
Machine gun	_____	_____	_____
Rocket	_____	_____	_____
Mortar	_____	_____	_____
Recoilless rifle	_____	_____	_____
Chemical (CS)	_____	_____	_____
Other ____________	_____	_____	_____

2. Ammunition Problems: Adequacy, capability, appropriateness) (Indicate comments or notations from battalion records.)

____

____

____

3. Fire Effectiveness: (Provide any comments or notations from battalion reports. Indicate if no comment made. Cover friendly and enemy.)

____

____

____

Additional Comments

____

____

____

____

AMMUNITION VARIABLES FORM #7

Figure 2. Example Data Collection Form

TABLE 5

Topics for Objective Data Collection

Form #1	Personnel - Key Staff and Commanders
Form #2	Personnel Variables
Form #3	Unit Experience
Form #4	Unit Variables: Friendly-Enemy Factors
Form #5	Equipment Variables - Initial
Form #6	Equipment Variables - Losses
Form #7	Ammunition Variables
Form #8	Supply and Evacuation Variables
Form #9	Unit Experience - Training
Form #10	Operations Variables - Support
Form #11	Unit Operations Variables
Form #12	Outcome Variables

Form #12, the last objective data form, is significant because it was used to collect, not variables potentially important in determining combat effectiveness, but traditional measures of effectiveness. Commonly applied quantitative criteria such as casualties given and taken, terrain gained or lost, speed of advance, and so forth, were collected here to allow later comparison with the judgmentally derived measures of combat effectiveness. The research hypothesis behind their collection was, of course, that they are unidimensional measures and would not, therefore, produce coherent relationships with predictor variables. The mission accomplishment judgments, on the other hand, are multidimensional and will produce associations with theoretically appealing predictor variables.

A NOTE ON DATA COLLECTION

The reader who has referred to Volume II to examine the coding forms will be aware that the data collection effort reported in this chapter extends to hundreds of variables. Ignoring the resulting analytical problems until later chapters, why seek so much information? The philosophy guiding the effort was to collect whatever was available, in as much detail as the research team felt would avoid the error of false precision. In this way, loss of available information was avoided. No extra effort was spent in returning to the files to obtain specific pieces of information, but no theoretically interesting variable went unexamined due to lack of research effort. The field of projecting combat effectiveness is virtually unexplored. Searches of the National Technical Information Service (NTIS) and the Defense Documentation Center (DDC) show less than a half dozen directly relevant research efforts during the past two decades. Failure to consider all possibly important factors would have been foolish.

As later chapters explain, many variables that were examined either proved to be so badly underreported in the available historical information as to make meaningful analyses impractical, or proved to have little association with measures of combat effectiveness. In these cases, the research effort has demonstrated the lack of relevant data or the absence of support for hypotheses involving these variables. Hence, the effort was well spent and contributes directly to the research literature. Moreover, as the research effort continues in future years, the number and variety of cases in the data base will grow. As this happens, it will become possible to carry out meaningful analyses involving more and more of these underreported variables.

CHAPTER 4. MEASURING COMBAT EFFECTIVENESS

INTRODUCTION

Probably the most important idea in this research project is that of using experienced officers and historically based case narratives to create a variable, "combat effectiveness." Without the success of this effort, the research results could not be considered "reliable" or "valid." That is, there would be no basis for assuming that the methodology employed in this study would produce the same results if repeated (reliability of the methodology), nor would there be reason to believe that the research findings accurately reflect the phenomenon they are supposed to reflect (validity of the project). If successful, the methodology provides a dependent variable -- a new measure of combat effectiveness. With that variable, it becomes possible to examine systematically any number of other factors that theoretically predict combat effectiveness to determine whether, in fact, they are associated with this phenomenon.

This chapter reports the data analyses and findings related to combat effectiveness. There is one basic hypothesis under analysis:

- The judgmental coding of combat effectiveness based on military mission accomplishment will form a coherent scale for the 22 engagements.

JUDGMENTAL MEASURES OF COMBAT EFFECTIVENESS

Questions Asked

The officer participants answered 13 questions about the effectiveness of unit performance for each engagement that they rated. As indicated in

Table 1, these questions focused directly on the "combat effectiveness" of the unit, but they differed in two important ways. First, the questions were phrased differently. Some merely asked about combat effectiveness, some defined it in terms of military mission accomplishment, others asked the raters to consider extraneous factors such as terrain, enemy resistance, actions of superior headquarters, and quality of support. The purpose was to identify which factors the officers were considering, though the goal was to find one question that focused on overall unit performance.

The questions also differed in the ranges of scales -- some were 5 points, some 10, and some 100. The purpose of this difference was to generate as much variance as possible in the data, without obtaining false precision. The full questionnaire is reproduced in Volume II, Appendix B of this report.

Data Inspection

Simple inspection of the data (Table 1) suggests several things. First, the size of the standard deviations and the distribution of scores on the two 100-point variables suggest false precision -- the officers did not reliably discriminate among engagements across the 100-point scales.

Second, units are rated worse (higher rankings = worse performance) when their performance is evaluated in isolation (Questions 2 and 3) than when performance is examined in light of other factors such as terrain, higher headquarters, resistance encountered, and support received (Questions 4-7). Those units that perform best may not have been outstanding, but rather made use of terrain, outstanding support, or planning by higher headquarters. An alternative inference is that the location of the isolated performance measures at the beginning of the questionnaire led to initially critical scores for them and less discrimination later. This inference was rejected because scores on 3B were higher than 1, and

TABLE 1

Combat Effectiveness Judgmental Questionnaire

Question Number	Abbreviation	Focus	Rating Scale	Mean	Standard Deviation
1	EFF/100	Combat effectiveness	0-100	79.6	19.4
2	EFF/10	Combat effectiveness, 10 typical engagements	1-10	3.1	2.0
3A	MIS/5	Military mission	1-5	2.0	1.2
3B	MIS/100	Military mission, percentage accomplished	0-100	80.8	23.1
3C	MIS/10	Military mission, 10 typical engagements	1-10	3.2	2.2
4A	R-T/5	Performance given resistance and terrain	1-5	1.7	0.8
4B	R-T/10	Performance given resistance and terrain, 10 typical engagements	1-10	2.8	2.0
5A	HHQ/5	Performance given actions of superior headquarters	1-5	1.6	0.8
5B	HHQ/10	Performance given actions of superior headquarters, 10 typical combat engagements	1-10	2.6	1.9
6A	SUP/5	Performance given adequacy of support	1-5	1.8	0.9
6B	SUP/10	Performance given adequacy of support, 10 typical combat engagements	1-10	2.8	1.8
7A	ALL/5	Overall performance in light of resistance, terrain, actions of superior headquarters, and adequacy of support	1-5	1.7	0.9
7B	ALL/10	Overall performance in light of resistance, terrain, actions of superior headquarters, and adequacy of support; 10 typical combat engagements	1-10	2.6	1.7

units were rated less effective on 3C than 2, 6B than 5B, and 6A and 7A than 5A. The first research finding is therefore,

- Units are rated less effective when overall "combat effectiveness" and "military mission accomplishment" are evaluated in terms of the unit in isolation. They are rated more effective when contextual factors and/or the performance of superior headquarters and/or support units is included in the analysis.

This can be read as reflecting common sense -- that an infantry battalion does not operate in a vacuum, but rather in a world of obstacles and constraints. More significantly, however, it also implies that <u>adaptability</u> is an important element in determining performance. The ability to react to different levels of resource availability, to take advantage of opportunities provided by terrain, and to react to enemy resistance would be expected, based on this finding, to be an important discriminator between successful and unsuccessful units.

Inspection of the data also showed that there was no essential difference between the 5-point scales and the 10-point scales. Since the 10-point scales would produce greater variance, and hence allow more sophisticated and effective data analysis, they were usually chosen for subsequent use.

Scores on Typical Questions

The scores for typical questions are shown in Table 2. Question 1 employs a 100-point scale; the others are 10-point scales. Higher scores on Question 1 are good. Low rankings on all other questions indicate excellent performance. The ID numbers refer to the 22 operations under analysis. The names of the engagements have been deliberately omitted to encourage the reader to think about the properties of the data set itself. Note that there is a great consistency across the 10-point scales for any given engagement and that many of the values assigned reflect excellent performance.

TABLE 2

Average Scores for Selected Questions

ID Number	1 EFF/100	2 EFF/10	3C MIS/10	4B R-T/10	5B HHQ/10	6B SUP/10	7B ALL/10	# of Responses
	Question Number and Abbreviations							
1	87	3	4	1	1	2	2	3
2	89	2	2	2	1	2	1	3
3	87	2	1	2	1	1.5	1	2
4	53	5	6	3	2	4	3	3
5	93	2	2	1.5	1	1.5	1	4
6	98	1.5	1	2	1.5	1.5	1.5	2
7	87	3	4	5	3.5	3	3	4
8	54	4	5	4	4	5	5	5
9	86	1	2	2	2	3	2	3
10	85	2	3	2	2	1	1	3
11	91	1	2	1	2	1	1	5
12	73	3.5	3	3.5	3	2	3	2
13	90	2	3	2.5	2.5	4	3	2
14	78	4	4	2	2	3	2	3
15	88	2.5	2.5	1.5	1.5	1.5	2	2
16	73	3	3	3	4	3	3	3
17	80	2	2	3	3	2	2.5	4
18	84	3	2	2	2	3	3	3
19	85	2	1	1.5	2	3	2.5	2
20	93	1	1	2	2	1	1	3
21	90	3	2.5	4	3.5	3	3	2
22	39	7	7	5	5.5	2.5	6	4

Factor Analyses

The key methodological issue related to this data set is whether or not it will form the basis for a coherent scale of combat effectiveness. A good scale is reliable in that a variety of similar efforts will produce similar measures. A variety of techniques can be applied to determine coherence of a data set. One of the most powerful is factor analysis. Basically, factor analysis is a technique developed to locate clusters of variables that have similar scores. Some scholars think of it as an effort to locate redundancy. It is often employed where a large number of similar variables are being examined. In this context, the technique will demonstrate that there is a set of variables which are so similar that any one of them can be used as a surrogate or "marker variable" for the set. This allows "data reduction" or simplification of an analytical problem by permitting the analyst to work with a smaller set of variables while ensuring that vital information has not been lost through the elimination of variables.

Factor analysis produces "factors" that are vectors, or paths, through the clusters in the multidimensional data set. It is vital to understand that factor analysis is quite capable of disaggregating a highly related data set into several components. The researcher, by setting the limits for the amount of redundancy or "communality" that is considered in defining a cluster, has some control over the level of discrimination, but there are accepted "default values" for these controls in different disciplines. The technique is also completely statistical -- it will detect theoretically nonsensical clusters as readily as meaningful ones. It is used to explore the "dimensionality" of a data set -- to measure its coherence and to locate its components. Frequently researchers have great difficulty in "naming" factors because they appear to be composed of items that correlate spuriously but do not form meaningful concepts.

The results of a factor analysis for the full set of 13 measures of combat effectiveness are shown in Table 3. Only a single factor, or cluster,

TABLE 3

Results of Factor Analysis
on Judgmental Combat Effectiveness

Question Number	Factor Loading	Estimated Communality
1	-.89	$\boxed{.91}$
2	$\boxed{.90}$	$\boxed{.92}$
3A	$\boxed{.90}$	.85
3B	-.81	.84
3C	$\boxed{.90}$	$\boxed{.91}$
4A	$\boxed{.90}$	.88
4B	(.76)	(.78)
5A	.84	.88
5B	.83	.89
6A	(.76)	.86
6B	.86	$\boxed{.92}$
7A	.88	.84
7B	.88	.80

Eigenvalue = 9.77

$\boxed{}$ = highest (over .90)

( ) = lowest (below .80)

was located. Its high eigenvalue (9.77) is very strong (usually any factor with an eigenvalue over 1.00 is considered worthy of attention).

Two types of information are given for the relationship between the scores from each question and the factor or vector that traces their redundancy. The factor loading indicates the extent to which the variable scores are associated with the values of the factor. The estimated communality focuses on the extent to which the variable participates in the redundancy being examined. These numbers are theoretically bounded between 0 and 1.00. Typically, a loading or communality of 0.30 or higher is seen as signalling a meaningful relationship between a variable and a factor. The values for this factor are very high. The blocks around values over 0.90 and parentheses around those below 0.80 call attention to the highest and lowest values. They have no statistical significance.

Factor I clearly finds the data on combat effectiveness to be extremely coherent. Over one-half of the variance in every question is associated with the vector (loading greater than 0.71), and all variables participate heavily in the communality. Hence, the second research finding,

- Judgmental measures of combat effectiveness form a coherent, unidimensional scale for the 22 engagements examined.

Marker Variable

Once a theoretically meaningful and statistically sound factor has been identified, there are two ways of proceeding. It is possible to generate "factor scores" for each case in the analysis. If these are determined, a new variable is created -- called Factor X -- representing the cluster of variables being examined. The problem with this procedure is that Factor X is neither replicable nor directly interpretable. It is not replicable because the addition of virtually any new data into the set will produce a different solution to the factor analysis, and hence new factor scores for all cases. It is not directly interpretable because it is a statistical artifact composed of positions of observations of reality, and not an observation of reality in itself. The alternative, selection of a marker variable, was chosen to allow for later expansion of the data set and to preserve direct capability to interpret results.

Returning briefly to Table 3, it is clear that Question 2, a 10-point ranking of combat effectiveness, is an excellent marker variable. It ties for highest factor loading and for highest estimated communality. One of its rivals, the 100-point question, was found to imply false precision and be somewhat unstable. The other, 3C, calls for evaluating military mission accomplishment, but also has the highest average scores, suggesting that the officer participants may have been more lenient in coding it than Question 2. The fact that these three questions were the leading variables is very encouraging since the objective of the questionnaire was to capture estimates of unit performance and ability to overcome obstacles. Higher loadings on the questions that focus on external factors (terrain, enemy, and so forth) would imply that the factor was related heavily to context of the situation. Hence, Question 2 was selected as a marker variable to represent the set of judgmental variables.

Data relating to the 22 engagements, based on Question 2 (10-point scale), are shown in Table 4. The first section of the table shows the scores within each era or type of combat. Note that there is variance present for the scores for all four groups. The mean values for the four groups are 2.65 (World War II), 2.7 (Korea), 2.6 (Vietnam), and 2.55 (Special Operations). Given the rather small n, these are remarkably close together. Hence, a research finding,

- No bias was detected in outcome coding for the different combat eras.

TABLE 4

Rankings of Operations:
Marker Variable, Question 2

A. Within Groups

ID Number	Short Title	Score
	World War II	
9	Iwo Jima, Suribachi	1
10	Iwo Jima, North	2
17	Guam II	2
19	Saipan	2
15	Okinawa, Oroku	2.5
1	Peleliu I	3
16	Guam I	3
21	Cape Gloucester	3
14	Okinawa, Motobu	4
4	Peleliu II	5
	Korea	
11	Yongdungpo	1
20	Yudam-ni Breakout	1
6	Inchon	1.5
18	Seoul	3
22	JAMESTOWN	7
	Vietnam	
2	STARLITE	2
3	DEWEY CANYON	2
5	Hue City	2
7	Khe Sanh I	3
8	Khe Sanh II	4
	Special Operations	
13	BLUEBAT (Lebanon)	2
12	Dominican Republic	3.5

Continued

Table 4
Operation Rankings
Continued

B. All Engagements

ID Number	Short Title	Score	Conflict
9	Iwo Jima, Suribachi	1	WWII
11	Yongdungpo	1	Korea
20	Yudam-ni Breakout	1	Korea
6	Inchon	1.5	Korea
2	STARLITE	2	RVN
3	DEWEY CANYON	2	RVN
5	Hue City	2	RVN
10	Iwo Jima, North	2	WWII
13	BLUEBAT (Lebanon)	2	Special
17	Guam II	2	WWII
19	Saipan	2	WWII
15	Okinawa, Oroku	2.5	WWII
1	Peleliu I	3	WWII
7	Khe Sanh I	3	RVN
16	Guam I	3	WWII
18	Seoul	3	Korea
21	Cape Gloucester	3	WWII
12	Dominican Republic	3.5	Special
8	Khe Sanh II	4	RVN
14	Okinawa, Motobu	4	WWII
4	Peleliu II	5	WWII
22	JAMESTOWN	7	Korea

The same information is presented in Section B of the table, except that engagements from different eras are combined. The listing for tied ranks is ordered by ID number assigned by the research team and has no significance.

The data do, however, show a distinct statistical problem -- low variance. Given that the officer participants were asked to rank the units on a 10-point scale, it is remarkable that the modal ranking is 2. The distribution of codings is shown in Figure 1. The distribution is skewed toward low rankings (excellent performance) and has few scores that indicate poor performance. This distribution of scores provides a severe handicap for statistical analysis -- lack of variance in the dependent variable. Hence, a finding,

- Further research should focus on improving the distribution of combat effectiveness scores, either by including new cases that involve poorer performance or by structuring more discriminant measurement instruments.

On the other hand, this set of findings provides considerable hope that it is possible to use expert judgment coding to determine thresholds for adequate performance. Because of the need for detailed data, a relatively short period of time available for the research, and the limited resources available to develop materials on each case, there was clearly a bias toward selecting engagements that had been previously researched in detail. It is hardly surprising to find that these include a bias toward good performance. One of the objectives of this research effort is to help formulate information about "how much is enough." Hence, finding a number of cases of good performance will, in the long run, be of great value.

Creating Variance

For purposes of the present study an effort was made to use the data set for producing a measure that would both validly reflect the judgmental

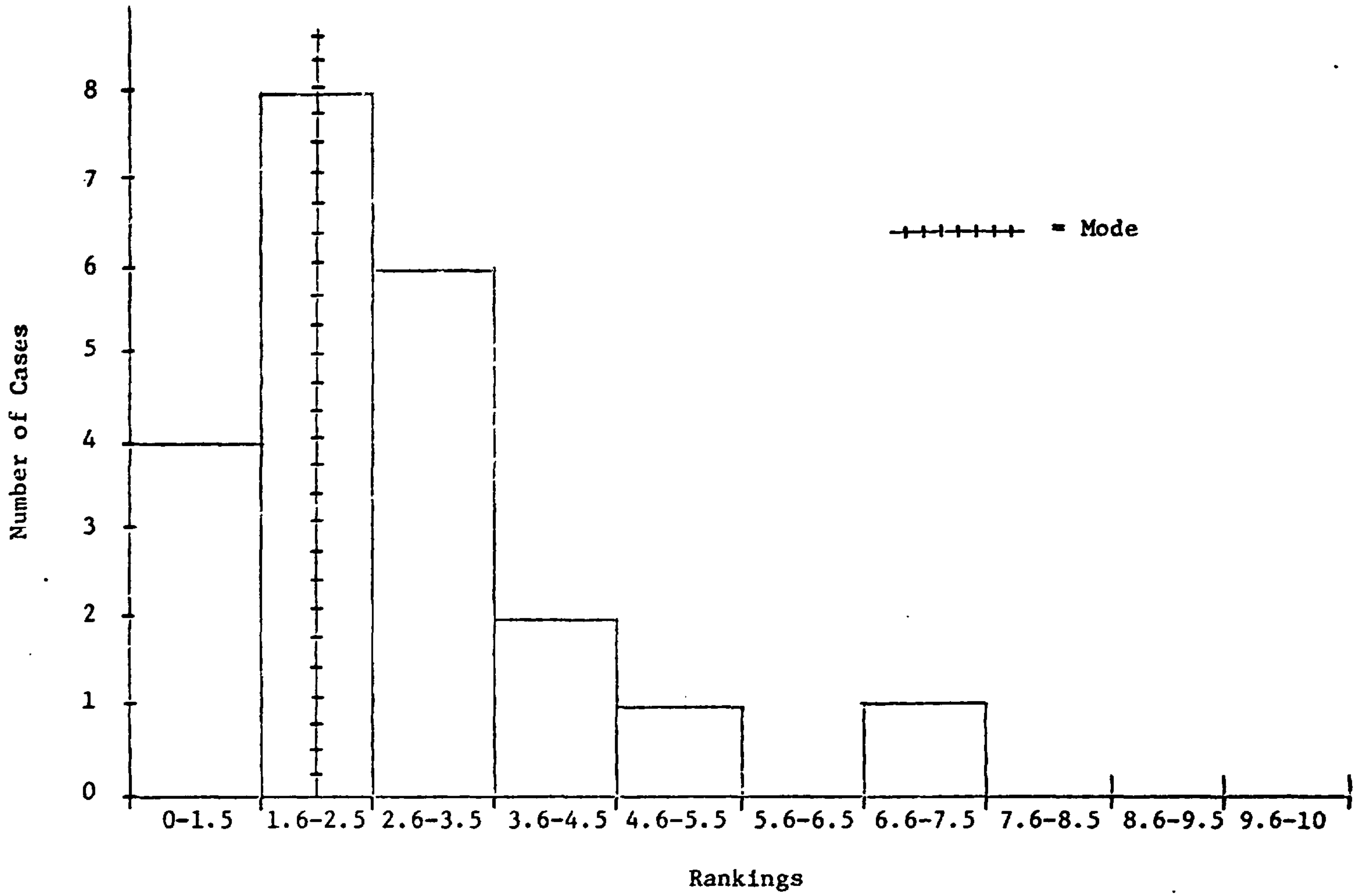

Figure 1. Distribution of Scores on the Marker Variable

data and have greater variance. The method selected was to create fairly rigorous, empirically based, discrimination points on each of the 13 judgmental questions. These points were labeled as distinguishing between "satisfactory" and "unsatisfactory" performance. They are close to the mean values for all variables, but are defined judgmentally by the research team.

This data transformation has important theoretical implications. First, it focuses the analysis at a comparative level -- it will explain which units are better than others. It will not indicate thresholds -- how much is enough. Second, it is a statistical artifact based on consistency of scores, not a directly interpretable variable.

The results of the transformation are shown in Table 5. Each engagement has been coded either 1 (satisfactory) or 0 (unsatisfactory) for each of the 13 judgmental measures of combat effectiveness. For 100-point variables, at least 85 points had to be assigned for the case to be rated satisfactory. For 10-point ranking scales, a ranking of less than 2.5 was satisfactory. For 5-point scales, rankings of less than 1.5 were coded as unsatisfactory. The right-hand column shows the percentage of unsatisfactory (0) scores recorded for each engagement.

These data were used to construct Table 6, which can be compared directly with Table 4. In terms of rank ordering, only the World War II cases show significant changes. Both Guam cases show loss of position within World War II, while Peleliu 1 moves up relative to the others. Given that this scale reflects data from all 13 questions, while Table 4 is based on a single question, and the complex transformation performed, these are relatively minor differences.

Examination of the scores across the entire data set does, however, reflect some bias based on combat groups. The bottom of the scale remains, as in Table 4, a mixture of engagements from different eras. At the top

TABLE 5

Satisfactory (1) Versus Unsatisfactory Combat Performance[a]

ID Number	1	2	3A	3B	3C	4A	4B	5A	5B	6A	6B	7A	7B	Percentage Unsatisfactory (0)
	Question Number													
1	1	0	0	0	0	1	1	1	1	0	1	1	1	38
2	1	1	1	1	1	1	1	1	1	1	1	1	1	0
3	1	1	1	1	1	1	1	1	1	1	1	1	1	0
4	0	0	0	0	0	0	0	0	1	0	0	0	0	92
5	1	1	1	1	1	1	1	1	1	1	1	1	1	0
6	1	1	1	1	1	1	1	1	1	1	1	1	1	0
7	1	0	0	1	0	0	0	0	0	0	0	0	0	85
8	0	0	0	0	0	0	0	0	0	0	0	0	0	100
9	1	1	1	0	1	1	1	1	1	0	0	1	1	23
10	1	1	0	0	0	1	1	1	1	1	1	1	1	23
11	1	1	1	1	1	1	1	1	1	1	1	1	1	0
12	0	0	0	0	0	0	0	0	0	0	1	0	0	92
13	1	1	0	1	0	0	0	0	0	0	0	1	0	69
14	0	0	0	0	0	0	1	1	1	0	0	1	1	62
15	1	0	0	1	0	0	1	0	1	1	1	0	1	46
16	0	0	0	0	0	0	0	0	0	0	0	0	0	100
17	0	1	1	1	1	0	0	0	0	1	1	0	0	54
18	0	0	0	1	1	1	1	1	1	0	0	0	0	54
19	1	1	1	1	1	0	1	0	1	0	0	0	0	46
20	1	1	1	1	1	1	1	1	1	1	1	1	1	0
21	1	0	0	1	0	1	0	0	0	0	0	0	0	77
22	0	0	0	0	0	0	0	0	0	0	0	0	0	100

[a] For 0-100 scales, satisfactory = 85 or greater
For 1-10 ranking, satisfactory = less than 2.5
For 1-5 scales, satisfactory = less than 1.5

TABLE 6

Rankings of Operations Based on Satisfactory/Unsatisfactory Performance

A. Within Groups

ID Number	Short Title	Score (Percentage Unsatisfactory)
	World War II	
9	Iwo Jima, Suribachi	23
10	Iwo Jima, North	23
1	Peleliu I	38
15	Okinawa, Oroku	46
19	Saipan	46
17	Guam II	54
14	Okinawa, Motobu	62
21	Cape Gloucester	77
4	Peleliu II	92
16	Guam I	100
	Korea	
6	Inchon	0
11	Yongdungpo	0
20	Yudam-ni Breakout	0
18	Seoul	54
22	JAMESTOWN	100
	Vietnam	
2	STARLITE	0
3	DEWEY CANYON	0
5	Hue City	0
7	Khe Sanh I	85
8	Khe Sanh II	100
	Special Operations	
13	BLUEBAT (Lebanon)	69
12	Dominican Republic	92

Continued

Table 6
Operation Rankings
Continued

B. Across Groups

ID Number	Short Title	Score (Percentage Unsatisfactory)	Conflict
2	STARLITE	0	RVN
3	DEWEY CANYON	0	RVN
5	Hue City	0	RVN
6	Inchon	0	Korea
11	Yongdungpo	0	Korea
20	Yudam-ni Breakout	0	Korea
9	Iwo Jima, Suribachi	23	WWII
10	Iwo Jima, North	23	WWII
1	Peleliu I	38	WWII
15	Okinawa, Oroku	46	WWII
19	Saipan	46	WWII
17	Guam II	54	WWII
18	Seoul	54	Korea
14	Okinawa, Motobu	62	WWII
13	BLUEBAT (Lebanon)	69	Special
21	Cape Gloucester	77	WWII
7	Khe Sanh I	85	RVN
12	Dominican Republic	92	Special
8	Khe Sanh II	100	RVN
16	Guam I	100	WWII
22	JAMESTOWN	100	Korea

end, however, no World War II cases escaped being ranked unsatisfactory on at least one of the 13 different questions. The scores for the list of Vietnam era cases are higher than they are on the overall combat effectiveness question reflected in Table 4. The Korean era scores retain their relative positions.

This suggests that Marine Corps infantry battalion performance in the World War II era was perhaps influenced more by forces beyond the unit's control -- higher headquarters, supporting fires, terrain, and enemy. This would be consistent with the finding noted earlier that the ratings on variables focusing on overall effectiveness (Questions 1, 2, 3A, 3B, and 3C) are higher than those that consider intervening variables. It is also consistent with the doctrinal requirement that battalions in Vietnam operate more independently, while those in World War II operate within an overall plan and structure controlled from above.

The technique also does what it is supposed to do -- it creates variance. Figure 2 shows the distribution of the transformed variable. A high score shows a high percentage of "unsatisfactory" ratings in comparison with the overall data set. The distribution of scores is very wide, and the mean value would be 44, near the center of the scale. Hence, as long as theoretical issues regarding which cases are best within this set are relevant, this new variable (percentage of unsatisfactory ratings) can be utilized.

COMPARISONS OF JUDGMENTAL DATA WITH HISTORICAL DATA

Purpose

It is a basic premise of this research effort that no good objective or historical measure of combat effectiveness has been identified in past research. Indeed, as noted earlier, very few efforts have been made to examine the topic. Therefore, the study presumes that it is worth the effort and expenditure to create one by systematically aggregating judgmental codings and analyzing them.

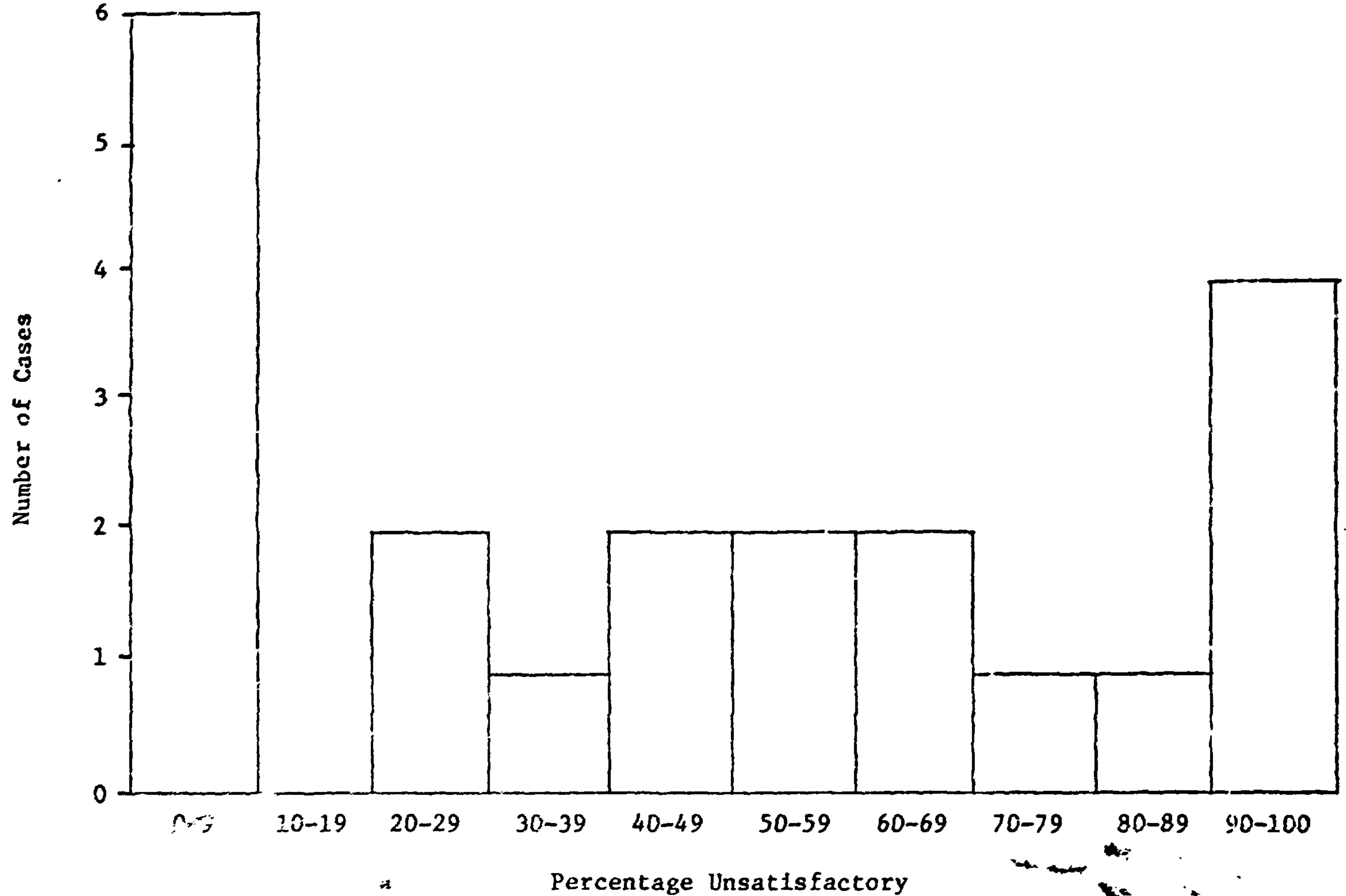

Figure 2. Distribution of Satisfactory/Unsatisfactory Performance Scores (Percentage Unsatisfactory)

To test this assumption, it was necessary to develop the historical or "objective" data relevant to the 22 engagements being studied and compare this information with the judgmental measures being generated.

The predictions of the research team were that the historical data would

- Be difficult to gather, and
- Would not correlate with the judgmental measures of combat effectiveness.

These hypotheses were based on the ideas that combat effectiveness is a multidimensional phenomenon in which objective measures applied to one engagement are inappropriate to others, and that precise reporting of objective data regarding performance is seldom carried out during periods of intense combat.

Variables Considered

The CACI research team coded a number of "objective" or historical measures of the outcomes of the 22 engagements. These measures were chosen to reflect traditional ways of evaluating unit effectiveness. The variables relate to a "yes/no" coding of mission accomplishment, movement of the forward edge of the battle area, occurrence of enemy breakthrough, coordination with friendly units, residual capabilities (personnel, arms, vehicles, fuel availability, ammunition levels, and effectiveness of fire by friendly units) (see Figure 3).

For most of these variables, one of two factors prevents any effective data analysis:

- The historical record provides clear, precise answers in so few cases that meaningful analysis is not possible (insufficient n), or

ENGAGEMENT NO. _____

Operation (Name, location, year) ______________________________

Phase ____________ Dates ________________ to ______________

Unit (Bn, Regt, Div) ______________________________________

OUTCOME MEASURES

1. Mission

_____ OBJ Achieved? (Yes, No)

_____ FEBA MVT _____ On time? (Show hours late or early)

2. Enemy Casualties

_____ No. POW's Taken

_____ Enemy KIA

3. Enemy Breakthrough (if it occurred) (Check one)

Blocked Partly Sealed Off Allowed

4. Coordination With Adjacent Units (Check one)

Maintained Lost

5. Interior Contact

Maintained Lost

6. Residual Capabilities (Summarize from other forms)

A. Personnel (From Form #13)

________ % KIA

________ % WIA

________ % MIA

________ Final % of Authorized Strength

______________________________________ Explain abnormality

____

OUTCOME MEASURES (Summary) Form #12 (2 pages)

Figure 3. Outcome Measures Coding Form

Continued

Figure 3
Outcome Measures Coding Form
Continued

ENGAGEMENT NO. _____

Operation (Name, location, year) ______________________________

Phase ____________ Dates ________________ to ______________

Unit (Bn, Regt, Div) ______________________________________

OUTCOME MEASURES (Page 2)

B. Arms - Loss Rate (From Form #6) (Check one)

_______ _______ _______
Subnormal Normal High

C. Vehicle Loss (From Form #6)

______________________________________ Explain abnormality

____

D. POL - Availability (From Form #8) (Check one)

_______ _______ _______
Subnormal Normal Above Normal

E. Ammo Expenditure (From Form #7)

______________________________________ Explain abnormality

____

F. Ammo Supply (From Form #7)

______________________________________ Explain abnormality

____

G. Fire Effectiveness (From Form #7)

______________________________________ Explain if rated "Fair"

____

____

OUTCOME MEASURES (Summary) Form #12 Page 2

Figure 3. Outcome Measures Coding Form

- The values for the variables are so similar for all cases that they cannot be analyzed statistically (insufficient variance).

For example, enemy casualties were obtained for fewer than 10 of the engagements under study. For the most part, even the enemy forces are uncertain of which U.S. battalions are responsible for which casualties. In the World War II cases, the greatest detail available was often enemy division level. Accurate assignment to specific battalions within the relatively narrow boundaries of time and space covered by a single engagement was seldom possible. Another example, movement of the forward edge of the battle area, was coded as not applicable in 15 of the 22 cases.

Friendly KIA, maintenance of interior contact, and ammunition expenditure were the only variables coded with sufficient frequency and variance to allow meaningful analysis. The maintenance of contact and ammunition expenditure variables are analyzed in Chapter 6 with the other historical data because they proved to be meaningful only at an ordinal level of measurement.

Friendly casualties, however, were compared with the judgmental estimates of combat effectiveness using Pearson product moment correlations in Table 7. The hypothesis under examination would be that

- High combat effectiveness correlates with low friendly casualties.

The research team expectation, based on the argument that combat effectiveness is a difficult, multidimensional phenomenon, was that the hypothesis would not be supported by the evidence. As Table 7 shows, there is virtually no evidence that the hypothesis is correct. Only one measure of effectiveness (mission accomplishment on a 10-point scale) shows a reasonably significant correlation with the judgmental combat effectiveness

TABLE 7

Bivariate Correlations Between Total Casualties and Judgmental Estimates of Combat Effectiveness

Judgmental Measures of Combat Effectiveness	Killed in Action		Wounded in Action		Total Marine Corps Casualties	
	r	s	r	s	r	s
EFF/10	.28	.15	.39	.07	.19	.21
MIS/10	.37	.09	.50	.03	.33	.07
R-T/10	.07	.40	.09	.38	-.06	.40
HHQ/10	-.08	.39	-.08	.39	-.04	.43
SUP/10	.27	.17	.18	.26	.05	.41
ALL/10	.07	.40	.14	.32	-.03	.44
	n = 15		n = 15		n = 21	

r = Pearson product moment correlation coefficient
s = Level of significance
n = Number of cases

measures, and it appears only in relation to those wounded in action. Hence, there are three research findings:

- "Objective" or historical measures of combat effectiveness are costly and time consuming to collect at the level of infantry battalions and do not vary widely across cases.
- Friendly casualties are not a good measure of combat effectiveness.
- Combat effectiveness is a multidimensional phenomenon that is not accurately reflected in any single objective number.

CHAPTER 5. JUDGMENTAL EVALUATION OF CRITICAL FACTORS

INTRODUCTION

In addition to coding each engagement for the "effectiveness" of the Marine Corps infantry battalion under study, officer participants were asked to indicate which of 49 different activities, factors, and functions they believed had contributed to the relative success or failure of the unit. The purpose of this query was to capture the subjective element of combat -- to develop a sense of what these experienced officers felt were the crucial aspects of unit performance. However, this information was solicited in the specific context of the combat engagement description just read. The goal was to identify, not theoretical knowledge, but the "signals" or "tip-offs" that the officers noticed and believed to determine combat outcomes.

The unit of analysis in this chapter is, therefore, the coding sheet of the individual officer participant for a particular narrative description. Although any variable can have up to 68 codings (the number of sheets completed), officers were always given the option of indicating that there was insufficient information to support an accurate coding. In fact, no variable was coded as "critical" on more than 47 forms. Moreover, the same engagement may appear in the data set more than once, since more than one officer may have coded it. However, no engagement could be counted more than three times.

The analyses were limited in two ways. First, only those factors discussed in the narrative descriptions could be considered for coding. Hence, holes in the historical records from which the narratives were written or gaps in the narratives themselves would result in gaps in the results. Second, the officer participants often chose to "no code" variables that were fully discussed in the narratives. However, the objective of the coding exercise was to learn what the officer participants believed to be the crucial factors, and it appears to have succeeded in achieving that goal.

DATA SET

Data Frequency

Of the original set of critical factors variables, nine were coded as "no information" so frequently that they had to be eliminated from the statistical analyses. They are listed in Table 1. Some (maintenance, special equipment, barrier materials, security on the move, and adequacy of protection) were very difficult to research and therefore were seldom detailed in the narratives. Others such as POL supplies, weather, and food and water were reported only where they were exceptional. Hence, unless they were utilized to select cases, the chances of collecting a meaningful number of cases were slight. Quality of enemy forces engaged was usually available, but only in general terms. Apparently the officer participants were frequently unwilling to infer that specific battalions encountered forces "typical" of those reported to be opposing U.S. forces in a general area.

Three variables, naval gunfire, preparatory air interdiction, and close air support had relatively small frequencies of coding (16-21) and could only be used cautiously in later data analysis.

Data Variance

A somewhat larger set of variables was eliminated from statistical analysis because they did not display sufficient variance to allow analysis. They are detailed in Table 2. Basically, these variables were coded at almost constant values across all cases being examined. The officer participants felt they had sufficient information to evaluate unit

TABLE 1

Problems With Data Frequency

A. Variables Eliminated Because of Insufficient Frequency of Coding

- Weather
- Quality of Enemy Forces Engaged
- Adequacy of Protection
- Security on the Move
- POL Supplies
- Food and Water
- Barrier Materials
- Special Equipment
- Maintenance

B. Variables With Low Frequencies (16-21)[a]

- Naval Gunfire
- Preparatory Air Interdiction
- Close Air Support

[a] Variables included in subsequent analyses were coded 31 or more times.

TABLE 2

Variables Eliminated Because of Insufficient Variance in Coding

- Training
- Morale at Outset
- Consistency of Plan With Principles of War
- Timeliness of Orders
- Subordinate Unit Coordination
- Use of Fire
- Morale During Combat
- Discipline
- Aggressiveness
- Initiative
- Resourcefulness

performance, but did not discriminate among the values assigned across cases. Of course, variables that do not vary will always intercorrelate highly with one another, but the results are meaningless. Moreover, since the research team was already aware that the dependent variables did have reasonable variance, it was clear that these variables would have little or no explanatory power. If excellent morale is associated with both terrible performance and excellent performance, then there is no evidence that the two phenomena are related. Hence, there is no point in conducting analyses using variables with no variance.

Variables Utilized

After screening for statistical problems, the 13 variables shown in Table 3 were selected for use in subsequent analyses. As a set they are remarkably diverse, running the functional gamut from command to logistics and including five different categories of supporting fires. All have standard deviations equal to at least 20 percent of their mean values. Moreover, they are drawn from the full set of locations in the coding form -- indicating that officers did not merely get tired and stop coding as they moved through the data-generation exercise.

FACTOR ANALYSIS

Rationale

As with the set of judgmental combat effectiveness measures, the research team utilized a variety of techniques to explore judgmental codings of critical factors, then settled on factor analysis as the best approach for understanding the data. In this case, however, the issue was how many different dimensions were produced by the critical factors coding. Data reduction (the elimination of variables) by identifying redundancy was again important. Unlike the combat effectiveness data, however, the research team now expected to find a number of different dimensions (or

TABLE 3

Critical Factors Variables Utilized

Short Title	Mean	Standard Deviation	Number of Codings
Quality of Information	4.25	1.62	44
Quality of Plan	4.36	1.51	47
Logistics Support	4.33	1.47	36
Awareness of Enemy Capabilities	4.29	1.45	42
Implementation of Principles of War	4.72	1.09	36
Maneuver During Action	4.54	1.09	46
Artillery Support	4.06	1.50	34
Naval Gunfire	4.57	1.36	21
Preparatory Air Interdiction	4.69	1.40	16
Close Air Support	4.71	1.65	17
Armor Support	4.06	1.55	42
Linkages to External Units or Commands	4.14	1.52	42
Communications	4.66	1.21	38

set of intercorrelated variables). Moreover, examination of these clusters of variables was expected and intended to produce insights into the thinking of the officer participants -- to find out what types of variables they tended to view as related and which types they viewed as relatively independent.

A three-factor solution was found that was both statistically sound and theoretically appealing. Table 4 reports the data on this solution. The eigenvalues and percentage of variance show the relative strength or power of each factor in the solution. The factor loadings opposite each variable indicate the approximate correlation between the variable and the vector that describes the factor. The communality indicates the extent to which the factor analysis "explains" the variance in each named variable. To help the reader, squares have been placed around the highest factor loading for each variable, regardless of the absolute number. Circles have been placed around all values greater than 0.30. These two symbols are employed for all values that have relatively important associations with any factor.

Naming (Interpreting) the Factors

One of the most difficult and crucial elements in factor analysis is deciding what the factors mean. Of course, any factor is a statistical artifact and means nothing in the real world. Frequently, factor analysis solutions are found that appear to be based on spurious correlations -- no apparent logic can theoretically link statistically associated elements. But when a factor is found to suggest association between theoretically meaningful variable clusters, great care must be taken to name or "label" the factor clearly to prevent misinterpretation later when the statistical artifact is forgotten and the research concepts are being utilized.

TABLE 4

Factor Analysis of 13 Judgmental Critical Factors Variables

Short Title	Factor Loadings			Communality	n
	Factor I	Factor II	Factor III		
Quality of Information	.89	.38	.06	.93	44
Quality of Plan	.54	.69	.11	.78	47
Logistics Support	.48	.33	.20	.38	36
Awareness of Enemy Capabilities	.74	.34	.20	.70	42
Implementation of Principles of War	.31	.75	.22	.71	36
Maneuver During Action	.33	.61	.24	.54	46
Artillery Support	.58	.15	.31	.45	34
Naval Gunfire	.26	.65	.54	.78	21
Preparatory Air Interdiction	.25	.85	.27	.87	16
Close Air Support	.18	.22	.96	.99	17
Armor Support	-.22	.31	.76	.72	42
Linkages to External Units or Commands	.48	.32	.38	.47	42
Communications	.47	.15	.70	.74	38
Eigenvalues	6.46	1.82	.79		
Percentage of Variance	71.20	20.00	8.70		
Suggested Interpretation	Coordination Functions	Planning Functions	Supporting Fires		

Factor I has been labeled "coordination functions." The "functions" part is easy to explain since virtually all the activities in the data set are functions or the results of functions carried out during combat. It is tempting to call this factor "intelligence" because of the high loadings and communalities for "quality of information" and "awareness of enemy capabilities." However, a whole host of other factors are associated with the factor including mundane ones such as logistics, communications, linkages to external units, and quality of planning. Moreover, artillery support is hardly an intelligence function. Yet, every activity associated with the factor requires or facilitates coordination, including the acquisition and dissemination of information to ensure awareness of enemy capabilities. Hence, the label "coordination functions."

The second factor has been named "planning functions." This is somewhat easier to explain since the two types of support that load on the factor (naval gunfire and preparatory air interdiction) are primarily the results of planning before an engagement begins. There is also an element of planning flexibility in the factor because "implementation of the principles of war" loads <u>during the engagement</u>, as does "maneuver" <u>during action</u>. Support elements are also reflected in logistics, armor, linkages, quality of information, and awareness of enemy capabilities.

The "supporting fires" factor is the easiest to name. All types of supporting fire load on it except preparatory air (but close air is the strongest variable). Of all other variables, only communications and linkages to external units are associated with the factor. It is reasonable to presume that these are necessary prerequisites to the effective use of supporting fires on the battlefield.

<u>Marker (Representative) Variables</u>

Selection of marker variables to represent these three clusters in subsequent analyses proved to be a difficult task. This selection is based on the combination of factor loading and communality. Good marker variables have a high loading on a single factor and high communality in the overall solution. A variable with these attributes has values very close to those represented by the vector that describes the factor. A less common criterion, a reasonable number of observations, is necessary in this case. If a variable with a low frequency in the data set is utilized, large numbers of cases must be dropped from subsequent analyses or artificial scores invented to compensate for the missing values. These procedures distort analytic results and narrow the utility of research findings.

Returning to Table 4, four variables are unlikely candidates for marker variables for any of the factors. Logistics support, maneuver during action, artillery support, and linkages to external units or commands all have relatively low communalities. Hence, they are not strongly associated with the solution and do not represent good surrogates for any of the factors. In subsequent analyses, these variables could be expected to behave independently of any of the clusters or "factors."

By contrast, both quality of information and awareness of enemy capabilities are good candidates as marker variables for Factor I (coordination functions). Quality of information is probably best because of its higher communality, but both variables also show some relationship to Factor II.

It is very difficult to choose a good marker variable for Factor II (planning functions). "Preparatory air interdiction" has the strongest measures of association but was only coded as crucial in 16 cases. "Implementation of the principles of war" is a good second, but the variable itself is a collection of ideas, difficult to measure precisely or define in objective terms. "Quality of plan" is promising, but shows a strong attachment to Factor I as well as Factor II.. "Naval gunfire" is also associated with Factor III and only has 21 observations.

Factor III (supporting fires) has an exceptional association with "close air support," but covers only 17 codings. Armor support, with a surprising 42 for frequency, is a good marker, and "communication" would make a fine alternative.

RESEARCH IMPLICATIONS

Despite problems resulting from the statistical properties of the judgmental critical factors data, two significant research findings emerge from its analysis:

- Experienced infantry officers identify a wide ranging, theoretically balanced set of factors as explanations of success and failure in combat.
- The judgmental data set for identifying critical factors influencing combat effectiveness is composed of three statistically defined factors. Substantively, these factors can be interpreted as
  - Coordination functions,
  - Planning functions, and
  - Supporting fires.

CHAPTER 6. ASSOCIATIONS BETWEEN COMBAT EFFECTIVENESS AND OTHER VARIABLES

INTRODUCTION

The research project had three major goals. The first was successfully completed when it became clear that judgmental codings of combat effectiveness form a clear, reliable scale (see Chapter 4). The second major methodological issue was whether it would be possible to find meaningful associations, which would make theoretical and empirical sense, between predictor or "independent" variables and measures of combat effectiveness. Analyses of judgmentally derived critical factors estimates reported in Chapter 5 indicate that this methodological question has been resolved positively.

The "bottom line" for the research effort is, however, whether it is possible to produce meaningful substantive results -- whether valid and reliable associations can be shown between combat effectiveness measures and sets of predictive variables. This chapter focuses on that bott line problem. After a discussion of the caveats that must be understood in interpreting the research findings, the chapter examines, in detail, the statistical associations between judgmentally derived critical factors variables and combat effectiveness of units. This section includes specification of relative weights, or priorities, among the 13 critical factors found to be associated with unit performance. The next section focuses on the historical, or "objective," data collected by the CACI research team.

CAVEATS -- LIMITATIONS OF THE FINDINGS

No research effort can have universal application. Indeed, well-d ned research is usually narrowly focused to reduce the number of confounding variables present, allow greater confidence that the findings are meaningful, and to increase understanding of the processes at work. An outstanding effort in a narrow area will often allow analysis of much broader issues by subsequently relaxing assumptions or broadening the data base. A broadly based research effort will usually create large data collection problems, obscure relationships by including too many dissimilar cases, and prove difficult to interpret.

The major substantive limitations of the analyses presented below can be summarized in two main points:

1. The data are historical, not focused on the future. Hence, characteristics of the future battlefield such as enemy antiair, mechanized forces, air capability, and electronic warfare are either underrepresented or not present at all in the data set.

2. The focus is on Marine Corps infantry battalions in offensive missions. Extension to other types of units, levels of command, services, or classes of mission can only be made with great caution and conscious discussion of the theoretical problems created.

From a statistical standpoint, the reader must understand four principal limitations on the interpretation of the analyses. First, the analyses are cross-sectional, not over time. This means that they are based on differences among the 22 engagements, or the 50 critical factors coding sheets, the 63 ratings of combat effectiveness, or the 101 comparative rankings of combat performance, not on repeated observations of a single unit. To be truly valid, the findings should be tested by following units over time -- the same unit with different levels of commanding officer experience should produce findings similar to those from several different units, each of which has different levels of commanding officer experience. Plans have been made for doing this in subsequent research efforts.

Second, the analyses are limited by the historical record available. Neither the narrative descriptions nor the objective data collected by the study team are complete or perfectly accurate. They were based on current archival materials, which are imperfect and incomplete. The presumption is made that there are no consistent biases in the data set so that errors of omission are part of "random error."

Third, the absence of findings regarding a variable or class of variable indicates that no evidence was found indicating that the variable was systematically associated with changes (variance) in levels of combat performance. Hence, the analysis can omit factors or functions that are important in determining combat outcomes, but which have adequate values in all or nearly all engagements under study. In other words, the findings are stated positively: There is an association between variable X and combat effectiveness; therefore, the evidence is consistent with the hypothesis that X influences combat effectiveness. The absence of such a finding means that there is no evidence of such a relationship in this research, but the problem may lie in either data availability or case selection.

Finally, and perhaps most important, are the issues of numbers of cases and variance in the dependent variable. Statistically, 22 cases is a very small number with which to work. The results of the analyses are remarkably strong in light of this problem, but it did present difficulties during the analytic phase of the research. It was made even more serious by a lack of variance in the measures of combat effectiveness (see Chapter 4). For example, the research team rated only 5 of the 22 units as having failed to accomplish their assigned military mission. The average ranking of combat effectiveness assigned by the officer participants was 2.5 on a scale of 10, and only 4 engagements had average rankings worse than 3.0. This relatively high level of accomplishment in the data set meant that the traditional analytic techniques, which are based on variance analysis, were of somewhat limited utility.

Having noted these problems, there are two further points to be made. Data analysis was accomplished and the findings make excellent empirical and theoretical sense. Nevertheless, readers and potential users should be aware of these caveats and limitations. Second, based on the success of the methodology, a second research effort is planned that will directly resolve many of these problems. This new study will include

- Analyses of unit performance from the 1973 war, where the "emerging threat" is better represented.
- Analyses of cases in which U.S. Marine Corps battalions encountered shock effects and/or surprise from enemy force levels, positions, weapons, or tactics. This should
  - Provide more cases of imperfect performance, thus increasing the variance in the dependent variable.
  - Allow projection of the attributes of units that are prepared to withstand shock and surprise, a key element in the early days of conflicts.
- Analyses of units over time as well as cross sectionally.
- Enlargement of the data set and collection of variables that are focused to resolve ambiguities and uncertainties resulting from the first research effort.

COMPARING JUDGMENTAL CRITICAL FACTORS WITH JUDGMENTAL COMBAT EFFECTIVENESS

Bivariate Correlations

The simple statistical association between the 13 individual variables judged to b important determinants of unit performance and the 10-point combat effectiveness scale (Question 2, overall combat effectiveness) are reported in Table 1. The coefficients reported are Pearson product moment correlations. The number of cases and levels of significance are presented to help the reader gauge their relative importance across the set under consideration. A variable coded as crucial in every case would have an n value of 50, a strong (nearly -1.00) negative correlation coefficient,

TABLE 1

Bivariate Correlations Between Judgmental Critical Factors and Judgmental Measures of Combat Effectiveness

Variable	Correlation With Combat Effectiveness	Number of Cases	Level of Significance	Percent of Variance Explained
Quality of Information	-.39	44	.005	15
Quality of Plan	-.47	47	.001	22
Logistics Support	-.29	36	.041	9
Awareness of Enemy Capabilities	-.25	42	.055	6
Implementation of Principles of War	-.85	36	.001	73
Maneuver During Action	-.66	46	.001	44
Artillery Support	-.30	34	.042	9
Naval Gunfire	-.52	21	.008	27
Preparatory Air Interdiction	-.68	16	.002	46
Close Air Support	-.59	17	.006	35
Armor Support	-.42	31	.009	18
Linkages to External Units or Commands	-.38	42	.007	14
Communications	-.51	38	.001	26

and a high level of significance. Negative coefficients are the expected direction because they indicate a low numerical ranking on combat effectiveness as associated with a high value for the critical factor. The percent of variance explained is calculated by squaring the correlation coefficient.

Several items are noteworthy in this table. First, with the possible exception of "awareness of enemy capabilities," every variable has a strong statistically significant association with the combat effectiveness measure. This is unusual given the small number of observations. Logistics and artillery support are surprisingly weak in comparison with other factors. This may be due to a lack of detail in the narratives that were used to judge importance. But the fact that they were coded 34 and 36 times, respectively, would indicate that the information was often there, but does not covary with the combat effectiveness data as strongly as the other variables.

"Quality of information" and "linkages to external units or commands" are somewhat stronger. "Quality of plan" is also moderately associated with the combat effectiveness measure. Of the remaining support variables, "armor support," relevant in only 31 of the 50 codings, has the lowest coefficient, with naval gunfire and close air support having rising association. "Communication" has a fairly strong relationship. The powerful associations, however, are with "maneuver during the action," "preparatory air interdiction" (on the relatively few occasions when it is present), and "implementation of principles of war." The research finding from a review of this table would be that

- Judgmental codings of critical factors associate strongly with judgmental codings of combat effectiveness.

Two further points about Table 1 merit attention. First, the percentage of variance explained clearly totals well over 100 percent for the set of variables. This reflects the fact that the independent variables are interrelated at significant levels. As a result, it is possible to find the same change in the dependent variable associated with changes in more than one independent variable.

To take a simple example, if an individual loses his job and gets a divorce on the same day and then goes on a drunken binge that night, we may argue that he is drinking because of the loss of his job, the divorce, or some combination of or interaction between the two. There is no logical or statistical way to sort out the effect of these elements without more information (for example, he has lost nine jobs before this and has never before taken a drink in his life) or a theoretical framework (money is less important than a happy family life) that orders the data for us. When this problem of intercorrelated independent variables occurs, it is called "multicolinearity."

The highest bivariate correlation with combat effectiveness is "implementation of the principles of war." Because this is a composite variable (includes several different ideas) Table 2 was examined to explore its relationships with other judgmentally derived critical factors. The relatively high correlations with the entire set indicate the strength of the multicolinearity problem. The 0.82 correlation with "maneuver during action" and 0.76 with "preparatory air interdiction" indicate that elimination of the variable from the set would not help to clarify the relationship between other variables and combat effectiveness. In the absence of the variable "implementation of the principles of war," either or both of these two strong correlates would "explain" most of the variance attributable to it, leaving little residual variance for the other variables to influence.

Factor Analysis

The objective of the analysis was, therefore, shifted to understanding how combat effectiveness is related to the set of critical factors.

TABLE 2

Bivariate Correlations of "Implementation of Principles of War" With Other Judgmental Critical Factors

Critical Factor	Correlation
Quality of Information	.56
Quality of Plan	.67
Logistics Support	.37
Awareness of Enemy Capabilities	.49
Maneuver During Action	.82
Artillery Support	.38
Naval Gunfire	.54
Preparatory Air Interdiction	.76
Close Air Support	.37
Armor Support	.42
Linkages to External Units or Commands	.50
Communications	.45

A factor analysis was performed to determine this set of relationships. This analysis was identical in every respect to the one reported on the critical factors data in Chapter 5, except that the 10-point combat effectiveness measure was included in the data set.

The goal of this analysis is to see (a) how the structure of the factors (or dimensionality) in the data set changes and (b) to see which sets of variables associate with the dependent variable by loading on the same factors with it. This uses the multicolinearity of the critical factors data to create dimensions, then allows the researcher to identify meaningful patterns in the data. It also permits subsequent data reduction (the elimination of variables) by identifying marker variables representing sets of variables. These marker variables then take on new meaning -- they represent not only the data selected on the variable but also the concept or idea captured by the factor (or dimension) that they mark.

The results of this factor analysis are shown in Table 3. Comparison of the factor structure (loadings and communalities) with that in Table 4, Chapter 5, shows that the structures are almost identical. (This is by no means surprising since all of the original data are repeated, and only the new effectiveness values are potential sources of change.) There are two real differences. First, naval gunfire loads more strongly now on Factor III than on Factor II, and a fairly large drop is shown in the loading of naval gunfire on Factor II. This helps to clarify the interpretation of Factor III, which can be labeled "supporting fires" with greater confidence.

The second shift is in the variable "quality of plan." Its loading on Factor II drops even more dramatically and, in the final analysis, is closely associated with Factor I. This raises the possibility that Factor II may be somewhat mislabeled. Perhaps the term "planning and command functions" would be more accurate than "planning functions" since the loadings for two variables previously seen as supporting the planning interpretation -- quality of plan and naval gunfire -- shift away from the factor when combat effectiveness is included.

TABLE 3

Factor Analysis of Judgmental Critical
Factors and Judgmental Combat Effectiveness

Factor Loadings

Short Title	Factor I Coordination Factors	Factor II Planning and Command Functions	Factor III Supporting Fires	Communality
Combat Effectiveness	-.13	$\boxed{-.76}$	.36	$\boxed{.73}$
Quality of Information	$\boxed{.88}$	.30	.09	$\boxed{.87}$
Quality of Plan	$\boxed{.66}$	.51	.16	$\boxed{.73}$
Logistics Support	$\boxed{.54}$	.23	.23	.40
Awareness of Enemy Capabilities	$\boxed{.81}$	.22	-.17	$\boxed{.74}$
Implementation of Principles of War	.31	$\boxed{.89}$	.18	$\boxed{.92}$
Maneuver During Action	.33	$\boxed{.71}$	.20	.65
Artillery Support	$\boxed{.58}$	.09	.33	.45
Naval Gunfire	.41	.45	$\boxed{.59}$	$\boxed{.72}$
Preparatory Air Interdiction	.41	$\boxed{.69}$	.32	$\boxed{.75}$
Close Air Support	.19	.20	$\boxed{.94}$	$\boxed{.95}$
Armor Support	-.20	.30	$\boxed{.77}$	$\boxed{.73}$
Linkages to External Units or Commands	$\boxed{.46}$	.32	.37	.45
Communications	.40	.20	$\boxed{.68}$	.65
Eigenvalues	6.98	1.83	.94	
Percentage of Variance	71.60	18.80	9.70	

Examination of the factor loadings and communality of the combat effectiveness measure is also revealing. First, the communality is quite high -- over one-half of the variance in combat effectiveness is associated with this solution of the factor analysis. Second, the loadings have the expected signs (good performance, a low ranking, is associated with high critical factors scores). Finally, Factor II (planning and command functions) has a strong association with combat effectiveness, Factor III (supporting fires) a moderate one, and Factor I (coordination functions) a weak but correct direction association. The findings based on the analyses would read

> Combat effectiveness is associated quite strongly with the effective execution of command and planning functions, moderately strong with the availability and use of supporting fires, and somewhat with coordination functions.

In addition, the stronger loading on Factors II and III indicates that the actions of a unit <u>after</u> the inception of combat are the vital elements. The <u>adaptability</u> of the unit -- ability to react to the environment in which it is operating -- appears to be the central predictor of effective performance.

Regression Analyses

To develop an algorithm by which successful unit performance might be estimated, regression analyses were performed using the 10-point combat effectiveness scale as the dependent variable and different combinations of the judgmental critical factors as the independent (or predictor) variables. As with the other analyses discussed in this volume, those actually reported are only a fraction of those carried out. The number of permutations and combinations by which the 13 critical factors variables might be used to project combat effectiveness is very large. Selected results are presented and discussed.

The regression shown in Table 4 was performed early in the analytical phase. It uses six independent variables, two representing each of the statistical factors found to be inherent in the critical factors data set. The purpose of the analysis was to see whether combinations of these marker variables would produce a "good" regression equation that would allow combat effectiveness to be projected from a small number of variables. The use of 6 variables, rather than the full set of 13, was based on (a) knowledge that the multicolinearity problem in the critical factors data would render a 13-variable equation meaningless, and (b) the clear pattern of dimensionality reflected in the factor analysis, indicating that marker variables would represent the overall data set very well. In addition, parsimony, one of the basic rules of data analysis, always drives researchers to seek the simplest available explanation.

Several pieces of information are needed to evaluate a regression equation. The ideal equation has a high "multiple R," which reflects the extent to which the variance in the dependent variable is "explained" by the independent variables, and consequently a high (approaching 1.00) percentage of variance explained (adjusted multiple R^2). It has a relatively small standard error and a large F statistic.

The key to evaluating a regression equation, however, lies in the relationships between each individual independent variable and the dependent variable. These are summarized by B, the standard error of B, and beta. B represents the slope of the regression line selected to represent the relationship between the dependent and independent variables. A large B means that the leverage of the independent variable is great -- a small change in the independent variable is associated with a large change in the dependent variable. The standard error of B is the boundary of confidence around the estimate of B. If the standard error of B is larger than B itself, it is possible that B is estimated in the wrong direction, or does not differ significantly from 0 (has no leverage at all). Beta refers to a normalized B. Because different variables have different

TABLE 4

Regression Analysis Using Marker Variables From Judgmental Critical Factors to Project Judgmental Combat Effectiveness

Short Title	Associated Factor	B	Standard Error of B	Beta	Bivariate Correlation
Quality of Information	I	0.001	0.29	0.001	-.39
Awareness of Enemy Capabilities	I	0.508	0.36	0.241	-.25
Quality of Plan	II	0.057	0.26	0.042	-.47
Implementation of the Principles of War	II	-1.958	0.34	-1.049	-.85
Armor Support	III	0.316	0.25	0.240	-.42
Communications	III	-0.410	0.25	-0.240	-.51

The dependent variable is overall combat effectiveness in comparison with 10 typical engagements.

Multiple R = 0.90

Adjusted Multiple R^2 (approximate percentage of variance explained) = 0.73

Standard Error = 1.05

F Statistic = 10.5

scales, B is not directly comparable across variables. When all independent variables are "normalized" or placed on equivalent scales, leverage can be compared. Hence, beta is the weighting function of a regression analysis. Strong betas are a sign of a "good" regression.

Applying these criteria to the analyses reported in Table 4, we see that the regression analysis is not "good." It does have a high multiple R and adjusted multiple R^2, a relatively small standard error, and reasonably good F statistic. However, only two of the B's are in the expected direction ("implementation of the principles of war" and "communications"). All should be negative, but the multicolinearity problem causes shifting in the n dimensional space of the solution and alters the signs. Moreover, both "quality of information" and "quality of plan" have standard errors of B large enough to imply that their slopes may not be significantly different from zero.

Statistically Best Regression

When analyses are performed to generate a statistically sound equation using regression, the three best variables turn out to be "implementation of the principles of war," "communication," and "quality of information." The results of this regression are shown in Table 5. Note that the multiple R is nearly as high as with the six-variable equation, and the adjusted multiple R^2 is higher (0.78). The standard error is marginally smaller, while the F statistic is much larger. These improvements are all related to reducing the number of variables and the level of multicolinearity among the independent variables.

The coefficients for individual variables are also better. Two of the three variables have the correct signs on their B's. The multicolinearity problem is still present however (note the correlation matrix at the bottom of the table) and results in a positive B and beta for "quality of information." All standard errors of B are smaller than their B's.

TABLE 5

Statistically Strongest Regression of Judgmental Critical Factors Projecting Judgmental Combat Effectiveness

Short Title	Associated Factor	B	Standard Error of B	Beta	Bivariate Correlation
Implementation of the Principles of War	Planning and Command (II)	-1.63	0.23	-0.87	-.85
Communications	Supporting Fires (III)	-0.38	0.19	-0.23	-.51
Quality of Information	Coordination (I)	+0.28	0.16	+0.22	-.39

The dependent variable is overall combat effectiveness in comparison with 10 typical engagements.

Multiple R = 0.88

Adjusted Multiple R^2 (approximate percentage of variance explained) = 0.78

Standard Error = 1.02

F Statistic = 27.6

Pearson product moment correlations between independent variables

	Implementation of the Principles of War	Communication	Quality of Information
Implementation of the Principles of War	1.00		
Communication	0.45	1.00	
Quality of Information	0.57	0.50	1.00

Substantively, the equation is easier to interpret if the concepts developed in the factor analysis are remembered. The "planning and command" factor dominates the analysis both because of its high bivariate correlation with the dependent variable and because of its strong beta coefficient. "Supporting fires" maintains a clear relationship to combat effectiveness, but has considerably less influence (smaller beta). The "coordination" factor remains the weakest of the three, playing only a marginal role in the "explanation" and having the wrong sign on its B.

Theoretically Most Interesting Regression

In analyzing a data set, it is often useful to utilize a technique known as "stepwise" multiple regression. Under this procedure the regression problem is solved in a series of steps, with one independent variable at a time entering the equation. This approach allows the researcher to understand two types of "interactive" effects. First, it is possible that two variables, acting together, can explain a dependent variable much better than either of them alone. Stepwise analyses will detect this. Second, each variable that is entered in a regression equation is viewed as "explaining" some portion of the variance in the dependent variable. Once explained, that same portion of variance is generally not available to be explained again. Hence, a variable with a relatively low bivariate correlation with the dependent variable may, if there is multicolinearity among those independent variables with higher bivariate relationships, add more to the explanation of variance than any of them.

This phenomenon shows up in Table 6, which reports a stepwise regression in which the criterion for entering the equation was increase in "adjusted multiple R^2." The first variable to enter is always that with the highest bivariate relationship to the dependent variable -- in this case the "implementation of principles of war." The second variable, "awareness of enemy capabilities," was a surprise. Although it had the lowest bivariate correlation in the critical factors data set, it proved to be the variable with the greatest additional contribution. Gratifyingly, a representative variable from the third factor, "communications," entered next.

TABLE 6

Theoretically Most Interesting Regression of Judgmental Critical Factors Projecting Judgmental Combat Effectiveness

Short Title	Associated Factor	B	Standard Error of B	Beta	Bivariate Correlation
Implementation of the Principles of War	Planning and Command (II)	-1.65	0.26	-0.88	-.85
Awareness of Enemy Capabilities	Coordination (I)	0.30	0.18	0.21	-.25
Communications	Supporting Fires (III)	-0.27	0.20	-0.16	-.51

The dependent variable is overall combat effectiveness in comparison with 10 typical engagements.

Multiple R = 0.88

Adjusted Multiple R^2 (approximate percentage of variance explained) = 0.75

Standard Error = 1.02

F Statistic = 21.5

Pearson product moment correlations between independent variables

	Implementation of the Principles of War	Awareness of Enemy Capability	Communications
Implementation of the Principles of War	1.00		
Awareness of Enemy Capabilities	.22	1.00	
Communications	.45	.49	1.00

The substantive interpretation of this finding is that the coordination function is a vital one in projecting combat effectiveness. Hence, while statistically never particularly powerful, this dimension cannot be ignored.

The evaluation of this regression analysis (Table 6) is very similar to the previous one. Multiple R, adjusted multiple R^2, and standard error are comparable, although F statistic is slightly lower. There is somewhat less multicolinearity present, but only two of the B's have the expected signs. The standard errors of the B's are acceptable. The three factors bear about the same relationships as measured by the beta coefficients.

<u>Conclusions Based on Associations Between Critical Factors Codings and Combat Effectiveness</u>

The findings based on these analyses can be stated at two different levels. At the highest levels of abstraction, they suggest that "move, shoot, and communicate" is not a bad mandate for infantry units in the U.S. Marine Corps. Indeed, there would be grounds for real concern if, at these higher levels of abstraction, the research did not produce findings consistent with the previous experience of senior officers and the general theories of offensive action that have been developed in the past.

There are, however, some distinctive findings at even higher levels of abstraction:

- Supporting fires, when they are involved in an action, are likely to be extremely important, but they are associated strongly with only one element of the infantry battalion's own activities -- communications.
- Planning, command, and coordination are very tightly intertwined. The unit that cannot carry out these activities simultaneously will almost certainly fail.

- Planning and command during an engagement dominate the associations with combat effectiveness. Supporting fires and coordination are each perhaps one-quarter[1] as important as this set of functions.

- Coordination functions are an important element in determining combat effectiveness, but neither theoretically nor statistically do they become significant unless there is high-quality planning and command.

Dropping down one level of abstraction, it is also possible to draw conclusions at the level of the 13 individual variables drawn from the critical factors data set. Table 7 shows a breakdown of these variables and approximate weightings or relative priorities among them based on the percentage of variance that each variable explains in combat effectiveness. Many of these findings are important.

- The single most important variable, "implementation of the principles of war," is a composite referring to actions that the unit executes on the battlefield after an engagement begins. Being prepared to execute and react is vital. Here again, adaptive behavior appears to be central.

- The single most important function for unit success is maneuver during the action.

- Nonorganic supporting fires -- preparatory air, naval gunfire, and close air support -- are absolutely vital when they are involved in an action. Units must be trained to use them effectively if they are to achieve combat effectiveness.

- Communications are the second most important specific function that an infantry battalion must perform well to operate effectively in combat. Units that communicate well also have a good record in use of supporting fires, although specific linkages to external units or commands do not show up as critical in themselves.

---

[1] Based on the relative size of beta coefficients in the regression analyses carried out using marker variables from the three major factors.

TABLE 7

Approximate Importance of Judgmental Critical Factors in Determining Combat Effectiveness

Variables Important in Most Engagements

Short Title	Associated Factor	Approximate Weight[a]
Implementation of the Principles of War	II	7
Maneuver During Action	II	4
Communications	III	3
Quality of Plan	I	2
Armor Support	III	2
Quality of Information	I	2

Crucial Variables That Are Coded Less Than One-Half the Time

Short Title	Associated Factor	Approximate Weight[a]
Preparatory Air Interdiction	II	5
Close Air Support	III	4
Naval Gunfire	III	3

Variables Coded as Important, But Not Strongly Associated With Combat Effectiveness

Short Title	Associated Factor	Approximate Weight[a]
Linkages to External Units or Commands	I	1
Logistics Support	I	1
Artillery Support	I	1
Awareness of Enemy Capabilities	I	0.5

[a] Based on bivariate (Pearson product moment) correlations with judgmental combat effectiveness in comparison with 10 typical engagements. 1-10: Percent of variance explained.

- Quality of planning and quality of information are important contributions to combat effectiveness. They are perhaps four times as important as awareness of enemy capabilities. This probably means (a) that there are not a large number of cases in the data set in which awareness of enemy capabilities was poor, and (b) that the information and planning functions depend on knowledge of the entire situation -- terrain, weather, enemy, disposition of own forces, and so forth -- rather than on knowledge of the enemy situation.

- Effective use of armor support is an important contributor in slightly over one-half of the cases analyzed. Emphasis on armor support in training would be an important element in preparing infantry battalions for combat.

- There is evidence that logistics support and artillery support have a positive impact on combat effectiveness, but there is little evidence that they have been frequent determinants of combat outcomes.

COMPARING HISTORICAL OR "OBJECTIVE" DATA WITH COMBAT EFFECTIVENESS

Introduction

Data were collected on a large number of historical or "objective" variables. The logic was that there were two classes of variables that were better collected by the CACI research team than by officer participants -- unambiguous data such as numbers of personnel, amounts of ammunition, and so forth, and judgmental variables, which could only be coded by personnel who had read detailed information on the engagements, rather than the relatively brief narrative descriptions used by the officer participants.

These variables are examined to determine whether and how they suggest explanations for combat effectiveness that vary from the judgmentally derived critical factors data. Analysis focuses on the set of 23 variables listed in Table 8. These were selected because they met the statistical criteria of the study -- data available on at least 10 of the

TABLE 8

Historical or "Objective" Variables
Selected for Analysis

States of Preparation

U.S. Unit Composition and Experience

Completion of Training Cycle

Regimental-Level Training

Division-Level Training

Rehearsal Prior to Engagement

Surprise by the United States

Levels of Resistance

Enemy Unit Composition and Experience

Intensity of Infantry Resistance

Intensity of Artillery Fire

Intensity of Mortar Fire

Presence of Enemy Armor

Surprise by Enemy

Preparation Level of Enemy Positions

Supporting Fires

Preparatory Artillery Fire

Preparatory Air Strikes

Artillery Support During Engagement

Air Support During Engagement

Artillery Ammunition Supply

Mortar Ammunition Supply

Other Variables

Supply/Delivery Problems

Internal Contact

Evacuation

Ammunition Expenditure

22 engagements and sufficient variance across the 22 cases that some statistical inference was at least theoretically possible. Some variables at least theoretically possible. Some variables were also removed because of redundancy. Lack of variance proved to be a major problem for a number of variables in the operating environment, while data availability was a problem with personnel type variables.

Analytic Technique

A variety of analytic approaches were tried with this data set, as with the others. The major problem was lack of a large number of cases. The "law of large numbers," which underlies most of the commonly used inferential statistics, operates well above 30 cases, marginally down to 15. The small number of negative combat outcomes was also a problem.

The technique selected was to construct contingency tables based on the association between the historical variables and "satisfactory" levels of combat effectiveness. "Satisfactory" was defined in terms of a ranking of lower than 2.5 on the 10-point combat effectiveness scale (see Chapter 4, Table 5). This provides 11 positive outcomes and 11 negatives for the data set. Comparison of the distribution of cases across the set of historical variables against this expected 50 percent probability of success was viewed as indicative of the importance of the variable.

To help avoid subtle errors of interpretation and to clarify the relationships, a form of "extreme case analysis" was also adopted. There were six cases on which no combat effectiveness question was coded as unsatisfactory. There were three cases on which no codings of satisfactory performance were encountered. Taken together, these nine cases form the "best" and "worst" of the levels of performance. Variables that are powerful predictors of combat effectiveness should show a strong ability to discriminate between these two groups.

Performance Given Varying Levels of Preparation

Table 9 reports the first set of analyses -- those comparing preparation of a U.S. Marine Corps battalion with its performance in combat situations. This variable cluster was chosen because of the need to identify correlates or predictors of combat effectiveness before the commitment of units to action.

The first variable considered is the composition and experience of the Marine Corps unit. Three codes were possible -- veterans, regulars, or mixed (composed partly of regulars and/or veterans, but heavily of replacements or reserves). The probability of successful performance appears to have some tendency to go in the opposite direction from the prediction -- mixed units performing better than regulars or veterans. The extreme case analysis is not inconsistent with this interpretation. A reasonable research finding would be that

- Historically, USMC battalions composed of relatively more replacements and reservists have performed above average in combat situations.

Two factors may interact to produce this finding. On the one hand, veteran or regular units may grow stale or cautious as they build up experience. This explanation will be directly examined in the next research phase when unit performance is analyzed over time. On the other hand, this finding clearly indicates that the level of training and preparation, which the Marine Corps has been able to provide to individual replacements, has been of high quality, so that units receiving "fillers" are able to perform quite effectively.

The second variable covered is the unit training cycle. Here emphasis is placed on a battalion's ability to carry out the full set of prescribed training requirements, which vary from one combat era to another. These units are concentrated on those having only limited unit training

TABLE 9

Likelihood of Satisfactory Performance
by USMC Battalions With Varying States of Preparation

Variable	States of Preparedness	All Coded Cases: Probability of Satisfactory Performance	n	Extreme Cases: Probability of Satisfactory Performance	n
Unit Composition and Experience	*Veterans*	46	11	60	5
		47	17	50	6
	Regulars	50	6	100	1
	Mixed Units	60	5	67	3
Unit Training Cycle	Completed	67	6	100	2
	Accelerated or Limited	50	14	80	5
Regimental-Level Training	Completed	36	11	0	1
	Not Completed	60	10	75	8
Division-Level Training	Completed	40	10	0	1
	Not Completed	64	11	75	8
Rehearsal Prior to Engagement	Completed	44	9	0	1
	Not Completed	54	13	75	8
U.S. Surprise	*Present*	58	12	71	7
	Absent	29	7	50	2

or that executed their unit training on an accelerated schedule. The finding is again weak, but supported by the extreme case analysis.

- Units that have completed their full cycle of training seem to perform better than those trained on an accelerated basis.

The next two variables, regimental- and division-level training, show relatively strong findings in an unexpected direction. Training at these levels is associated negatively with combat performance. Indeed, six of the eight units that were rated satisfactory on every effectiveness question had not been trained at this level.

- Division and regimental training is negatively associated with successful combat performance, perhaps reflecting either overtraining of units or a failure to concentrate on the basics within the unit.

Since over 40 percent of the cases examined are taken from World War II, when much division and regimental training deals with the physical problems of amphibious landings, this finding may be somewhat suspect.

It must also be remembered that the operations utilized as cases in these analyses were chosen in part because the battalions were operating independently enough to identify the relevant historical materials. The nature of these operations may be such that higher-level training is not associated with success. These observations do not explain, however, why rehearsals are not associated with better outcomes.

- Rehearsals are not associated with above average performance by infantry battalions.

Finally, the Marine Corps battalions did perform better when they were able to stage a surprise. The large number of surprises is, of course, indicative of the selection of offensive operations for analysis.

- Surprise does tend to increase the probability of combat effectiveness.

Generally speaking, then, mixed units that have completed full unit training cycles and achieved some surprise on the field of action are most likely to perform effectively. At a minimum, there is no evidence that regimental-level training, division-level training, or rehearsals will assist a battalion in performing effectively, and these may be negative factors if the small data set reviewed here is typical of all engagements.

Combat Effectiveness in the Face of Different Levels of Resistance

The second cluster of historical variables to be examined deals with the enemy forces encountered (Table 10). One way of assessing U.S. Marine Corps battalions' needs is to see how they have performed in the face of different types of adversity.

The first variable, performance given the composition and experience of enemy forces, is a classic example of a "validating" finding. That is, the research reproduces something known, therefore greater faith can be placed in its other findings.

- U.S. Marine Corps infantry battalions are most effective against mixed and irregular enemy forces and somewhat less effective against veteran and regular units.

When compared with level of infantry resistance, infantry battalions show great strength -- the greater the enemy resistance, the higher the tendency to succeed. However, the next three variables show that, given heavy enemy artillery, mortar, and/or armored support, Marine Corps units perform less effectively. This is not surprising since the historical mission of these battalions, and their composition, is light infantry. It is however important.

TABLE 10

Likelihood of Satisfactory Performance by USMC Battalions Facing Various Levels of Resistance

Variable	States of Resistance	All Coded Cases: Probability of Satisfactory Performance		n		Extreme Cases: Probability of Satisfactory Performance		n	
Enemy Composition and Experience	Veterans	50	45	6	11	50	40	2	5
	Regulars	40		5		33		3	
	Mixed	57	54	7	11	100	100	3	4
	Irregulars	50		4		100		1	
Intensity of Infantry Resistance	High	67		6		67		6	
	Low	33		15		67		3	
Intensity of Artillery Fire	High	44		9		33		3	
	Low or Absent	50		12		80		5	
Intensity of Mortar Fire	High	46		11		50		4	
	Low or Absent	55		11		80		5	
Presence of Enemy Armor	Present	40		5		50		2	
	Absent	53		17		71		7	
Surprise by Enemy	Present	42		12		71		7	
	Absent	56		9		50		2	
Enemy Positions	Prepared	50		18		63		8	
	Hasty	33		3		--		0	

- Marine Corps infantry battalions tend to succeed when enemy infantry resistance is high. This may also imply that the enemy is failing to maneuver and allowing Marine Corps fire and maneuver to destroy them.
- Intense enemy artillery, intense enemy mortar fire, and the presence of enemy armor tend to reduce the effectiveness of Marine Corps infantry battalions.

A high priority must be given to preparing units to deal with the more lethal weapons possessed by enemy forces. Another interesting pair of findings is that neither surprise by the enemy nor prepared positions appear to offer any major advantage against Marine Corp infantry units.

- Marine Corps infantry battalions handle both prepared positions and tactical surprise situations well. These factors do not show a tendency to alter combat outcomes.

<u>Combat Effectiveness Given Varying Levels of Supporting Fires</u>

Table 11 reports data relevant to the third cluster of variables, supporting fires. This is particularly important because it focuses directly on one of the sets of ideas (statistical clusters or "factors") found important in the critical factors data set.

Six variables are available. Preparatory artillery fire and preparatory air deal with supporting fires before an engagement. Artillery and air support during an engagement are assessed separately and in conjunction with the relative availability of artillery and mortar ammunition.

Preparatory artillery fire shows a clear, "wrong way" association. The probability of satisfactory combat performance is clearly higher in situations where little or no artillery preparation occurs. This is true both in the 19 cases of coding and the 8 relevant extreme cases. This finding, one of the clearest in the data set, probably reflects the fact that

TABLE 11

Likelihood of Satisfactory Performance
by USMC Battalions Given Varying Supporting Fires

		All Coded Cases		Extreme Cases	
Variable	Levels of Support	Probability of Satisfactory Performance	n	Probability of Satisfactory Performance	n
Preparatory Artillery Fire	Moderate/heavy	40	10	33	3
	None/light	67	9	80	5
Preparatory Air Strikes	Heavy	33	3	50	2
		57	7	67	3
	Light/moderate	75	4	100	1
	Not made	55	11	60	5
Artillery Support During Engagement	Heavy	50	8	50	4
	Moderate	63	8	75	4
Air Support During an Engagement	Used	53	15	75	8
	Not used	40	5	0	1
Artillery Ammunition	Sufficient	50	4	67	3
	Marginal	57	7	50	2
Mortar Ammunition	Sufficient	54	13	75	8
	Marginal	80	5	—	—

moderate and heavy artillery preparations are usually carried out in anticipation of difficult combat. It is clear, however, that artillery preparations do not, by themselves, create a favorable combat situation.

- Preparatory artillery is used heavily in difficult situations, but does not, in itself, project improved combat effectiveness.

Preparatory air does show some impact on level of combat effectiveness. There is a small, predicted direction tendency for satisfactory performance to be more likely when some preparatory air is used. Light and moderate use of preparatory air does show association with higher probability of success. Heavy use of air, however, has a lower probability of success than either moderate, light, or no air preparation at all. Here, again, it appears that the heavy use of support in the preparation phase is confined to those cases where difficult combat is expected. When used, heavy air preparations do not, in themselves, create higher probability of satisfactory performance.

- Heavy preparatory air is used in difficult situations, but does not, in itself, project improved combat effectiveness.

- The presence of air preparations does have a small tendency to increase the probability of satisfactory performance.

- Selective use of preparatory air in moderate amounts has a tendency to produce increased combat effectiveness.

Analysis of the impact of artillery support during air engagements is complicated by the absence of cases in which little or no artillery was employed by the United States. Hence, analysis focuses on a truncated distribution. The heavy use of artillery is associated with difficult combat situations -- cases where the probability of successful performance is somewhat lower than normal. As with preparatory air, moderate artillery support during an engagement is associated with somewhat more success than heavy support.

- Artillery support during an engagement is most intense in difficult situations, but does not, in itself lead to improved combat effectiveness.

This finding, coupled with the earlier ones on preparatory artillery, effectiveness of enemy artillery fire, and the findings on crucial factors that artillery was weighted less than other forms of support suggest another finding.

- The employment of artillery is an area where considerable room for improvement in doctrine and execution may exist.

Air support during an engagement maintains its strong positive relationship to combat effectiveness. In the world of light infantry, close air support is capable of dominating the battlefield and clearly has done so in the engagements coded here.

- During an engagement close air support is the most effective type of fire support available to Marine Corps infantry battalions.

Ammunition supplies for supporting arms were also examined. The two-edged sword of intensity of combat and need for fire as a predictor of both marginal supply and low combat effectiveness (higher probability of not accomplishing a mission) is clearly present. Taken in conjunction with the findings about logistics in the critical factors data, it suggests that

- Supply systems are not generally capable of winning an engagement, but sufficient performance in this area is necessary to avoid negative outcomes.

Other Historical Factors

Three other historical variables produced sufficient variance to warrant examination (see Table 12). Overall presence/absence of supply problems was one. No strong relationship is revealed, supporting the argument that adequate supply may be necessary, but is unlikely to be a sufficient cause of combat success or failure.

The loss of contact among the components of a battalion (companies or platoons) has a clear relationship with mission accomplishment. This is consistent with the coordination function identified in the critical factors data.

- Maintenance of internal contact is essential in achieving successful performance.

Finally, ammunition expenditure has a negative association with positive outcomes. Again, difficult combat tends to consume ammunition, and light combat tends not to do so.

Conclusions Based on Historical Data

The historical data viewed in isolation provide fewer insights than the critical factors data. There are, however, some new findings. It is important to remember that these findings are based on a comparative standard of combat effectiveness, rather than the absolute one used with the critical factors data. That is, the <u>best</u> cases of performance are being compared with the worst, while "adequate" performance is not measured by the scale.

- Completion of a full cycle of unit training before commitment to a combat environment increases the probability of effective performance.

TABLE 12

Likelihood of Satisfactory Performance
by USMC Battalions Given Various Internal Problems

		All Coded Cases		Extreme Cases	
Variable	States	Probability of Satisfactory Performance	n	Probability of Satisfactory Performance	n
Supply Delivery Problems	Encountered	53	15	67	6
	Not Encountered	43	7	67	3
Internal Contact	Maintained	57	14	80	5
	Lost	38	8	50	4
Ammunition Expenditure	Low/Normal	56	9	100	3
	High	40	10	60	5

- Air support during an engagement increases the probability of effective combat performance of infantry offensive missions.

- Loss of internal contact among the components of an infantry battalion decreases the probability of satisfactory combat performance.

- U.S. Marine Corps infantry battalions have, in the cases studied here, performed less effectively when facing intense enemy artillery, intense enemy mortar fire, and enemy armor than when those factors were absent.

- Intense enemy infantry resistance has led to increased combat effectiveness by U.S. Marine Corps infantry battalions, perhaps by fixing the enemy in position.

- Marine Corps infantry battalions have performed more effectively against mixed and irregular enemy forces than against veterans and regulars.

- With the element of surprise on their side, U.S. Marine Corps battalions have had an increased probability of success, while they have generally been able to neutralize enemy tactical surprise situations.

- U.S. Marine Corps battalions composed of a mixture of regulars and veterans with replacements and reservists have performed as well as or better than regular and veteran units without reservists or replacements.

- Neither regimental-level training, division-level training, nor rehearsals for the specific engagements show a positive association with effective combat performance. The data suggest that they may detract from probability of satisfactory performance, perhaps by distracting the unit from more fundamental training.

- Prepared enemy positions have not caused lower probabilities of infantry performance. Like the intensity of infantry fighting, they may fix the enemy in a position where it can be destroyed by fire and maneuver.

- Artillery support during an engagement does not increase the probability of successful performance by a unit. Support is most intense during difficult combat. There appears to be room for improvement in doctrine and employment of artillery.

- Artillery preparations are negatively associated with effective performance. This may reflect the fact that difficult offensive missions are often preceded by heavy preparatory fires. However, these fires are not effective enough to bring up the probability of success to an equal level with other engagements.

- Moderate and light air preparations are likely to increase the probability of satisfactory effectiveness over those of engagements where no preparatory air is involved. However, heavy air preparations, associated by definition with difficult combat, do not in themselves increase the chances of successful combat.

- Ammunition expenditure, artillery ammunition availability, and mortar ammunition availability are not found to be either important aids or hindrances to effective combat performance, although consumption is higher during intense combat.

- Supply and delivery of supplies are not found to be either a major problem or a major determinant of combat effectiveness.

CHAPTER 7. EVALUATING MARINE CORPS INFANTRY BATTALIONS ENGAGED IN EXERCISES

INTRODUCTION

One objective of this research effort was to permit evaluation of a U.S. Marine Corps infantry battalion's potential for effective combat performance before it is deployed into hostile or potentially hostile environments. This can involve a variety of criteria -- types of weapons to be supplied, training to be emphasized, unit composition, and so forth. One of the best evaluation opportunities is, of course, the field exercise. In exercise situations units must "put it all together" -- move, shoot, and communicate. Hence, performance evaluations of exercises have been a major part of estimating unit preparedness for many, many years.

This chapter focuses on the question of how to use exercises and exercise generated data to evaluate a unit, the Marine Corps infantry battalion, to project its level of performance (mission accomplishment) in a hostile environment. The discussion is divided into several sections. First, a philosophy or approach to the execution of exercises is elaborated and reviewed. Second, existing systems of evaluation, including those developed by specific Marine divisions, are discussed, with emphasis on the newer systems. A third section suggests a logic for evaluation structure, based on ideas from the area of experimental design. The fourth discusses data to be utilized in evaluations. Finally, the need and opportunity for establishing a systematic data base relating to exercises are reviewed.

PHILOSOPHY OF THE APPROACH

Formally speaking, there are two approaches to evaluating the readiness of a unit in the U.S. Armed Forces:

1. "Readiness indices" reported by each unit, and

2. Exercises carried out either as part of training cycles or prior to deployment.

The system currently used to estimate the status of military forces focuses on unit readiness. Readiness is indicated by a number of unit "attributes" that are measured either objectively or subjectively. Hence, completion of required training hours, availability of crucial items of equipment, condition of equipment, estimates of unit morale, and a myriad of other tangible and intangible variables are considered in estimating a unit's level of readiness. Unfortunately, this system has a number of disadvantages. For example,

- Some of the measures are sterile and viewed as meaningless by many commanders.
- The intangible variables are estimated subjectively and wide disagreements can occur among qualified commanders.
- The system lacks historical validation.
- The relationship among the components of "readiness" is unclear, so a commander is given relatively little guidance as to priorities in upgrading unit capabilities.
- The system is relatively insensitive to improvement -- when a unit is deemed "ready," no upgrading is possible, regardless of new equipment or training.

Two criticisms of the system are particularly important. First, the readiness system is a "report card" not only on the unit commander but on every officer in the chain of command. To report a unit "not ready," or not up to its assigned level of readiness, is potentially damaging to a number of careers. The system generates strong downward pressure that is intended to reinforce the desires of unit commanders to perform well and produce combat-ready organizations. It can, however, lead to dysfunctional results since there is considerable incentive to report marginal areas as "ready." Second, there is a need for broad coverage of areas, thus excluding concentration on key areas. This type of structure will often slow the development of a unit by forcing "ticket punching."

By contrast, exercises of various sorts (tactical tests, training tests, and so forth) force units to perform against specific missions and provide opportunities for identifying key areas where additional training, planning, or types of equipment are needed. Almost by definition they avoid some of the problems with the readiness system.

- The exercise is dynamic and designed to test the capacity of the unit in the field. Hence, it has great potential to involve and motivate unit personnel.

- The intangible factors can be seen in context, especially the way they influence the unit's capability to accomplish its mission.

- Specific problem areas can be identified, and their impact on overall unit performance can be seen by the unit. Hence, priorities can be established for corrective actions.

- There are always areas where improvements can be made, and even a good unit can identify areas for future emphasis.

The exercises do, however, share the problem of being a "report card" for unit commanders. Downward pressure for excellent performance exists, and poor performance by a unit impacts negatively on the reputation of higher commands.

Exercises differ from the readiness system, however, in that they are learning experiences as well as opportunities for evaluation. These two factors interact to produce a general understanding that detailed findings about exercises will not be reported up the chain of command. Feedback at the unit level from umpires and exercise control personnel is frank (and often blunt), but the unit's ability to learn from the exercise is stressed in formal reports.

There is an opportunity imbedded in this system for developing a different way of evaluating units. The following discussion is predicated on a desire to use that opportunity. Basically the argument is as follows:

- Exercises are opportunities for both learning and evaluating.
- The CACI, Inc.-Federal research reported in this volume indicates that the "adaptive behavior" of a U.S. Marine Corps infantry battalion is the most important component of combat effectiveness.
- Learning is a form of adaptive behavior.
- CACI researchers found that it is possible to project combat effectiveness based on a relatively small number of unit functions.
- Therefore, it is both fair and wise to collect data relevant to absolute levels of performance on those indicators <u>and</u> on the learning (adaptive) behavior of the unit over the course of the exercise.

In other words, it is possible to evaluate unit potential for effective combat performance both on the basis of absolute performance and on improvement rates observed over time. This approach will not completely eliminate the problem of "report cards." It would, however,

- Force the use of more objective indicators, making identification of marginal areas of performance easier,
- Provide an opportunity to evaluate the unit for adaptability, a crucial area currently not examined, and
- Reduce the incentive for underreporting exercises both by creating more objective data and allowing demonstration of the learning curve for units during the exercise.

EXISTING EVALUATION SYSTEMS

While the research team was familiar, by experience, with a variety of different evaluation systems for unit readiness and exercise evaluation, a brief review of typical systems and measurement techniques was undertaken to ensure that new approaches were understood and no unnecessary effort was expended in "reinventing the wheel."

Research efforts carried out by the Naval Personnel Research and Development Center (NPRDC) provided a good background for this effort. As part of an ongoing project, that office had conducted a survey of research related to unit performance effectiveness measures (NPRDC, 1974). Virtually no historically validated systems were found in their survey. One major effort (McQuie, et al., 1969, 1968), entitled "Multivariate Analysis of Combat," was located that used historical data. This effort focused exclusively on World War II data and used the division as the unit of analyses. U.S. success was a judgmental variable coded by military historians. Focus was placed not on projecting success or failure but on comprehending the relationship between fire and maneuver. A factor analysis, for which incomplete statistical data are reported, revealed

- One factor with a 0.70 loading on U.S. success which, of 29 other variables, shows relationships only with mode of U.S. operations (attack/defend) (-0.47), mode of enemy operation (+0.55), and number of enemy medium tanks (-0.44). None of these are the highest loadings for their variables.

- Another factor with a loading of 0.49 for U.S. success has very strong loadings (0.86-0.90) for numbers of U.S. personnel, small arms, crew-served weapons, and mortars, and moderate loadings (0.41-0.52) for U.S. mode of operations, U.S. medium tanks, enemy mode of operations, enemy medium tanks, and enemy casualties (McQuie, et al., 1969: 66).

None of these findings is inconsistent with either common sense or the CACI research reported in this volume. However, it seems surprising that only medium tanks, of all possible variables representing maneuver, are associated with U.S. success.

The research team also examined readiness indices for ground unit training utilized by the Second Marine Division (2nd Marine Division, no date) to see how the readiness categories were converted into training requirements. Inspector General (IG) inspection forms and other standard evaluation sheets were also reviewed.

Documents relating to exercise evaluation were also obtained from Headquarters, U.S. Marine Corps. Among those examined were

- Standard Operating Procedures for Tactical Training Test, 2nd Battalion, 5th Marines, 1st Marine Division (1st Marine Division, 1977),
- Exercise Evaluation Sheets, 1st Marine Division (1st Marine Division, no date),
- BLT Tactical Evaluation Operation Order, Regimental Order 5041.2, 4th Marines, 3rd Marine Division, 29 March 1975 (3rd Marine Division, 1975),
- Predeployment Company Tac Test, 2nd Battalion, 9th Marines, 3rd Marine Division, 28 June 1966 (3rd Marine Division, 1976a), and
- Company/Platoon Tactical Testing, 3rd Battalion, 9th Marines, 3rd Marine Division, 9 March 1976 (3rd Marine Division, 1976b).

These systems are based on several assumptions. First, there are a large number of specific functions that units must execute in order to succeed. Second, umpires physically present in the exercise will be able to provide information on whether or not the unit performed adequately. Third, an exercise should be evaluated in terms of the entire package of activities. Most systems also include weighting, that is, some activities are viewed as more important than others. For these systems, a unit's total score is computed by multiplying the weights times the scores assigned. A few of these evaluation systems provide standards of measurement (0-25 is unsatisfactory, 26-50 is poor...) to be applied by the umpires.

Evaluation systems are good training tools. They force the units to deal with a variety of activities and cause exercise scenarios to be broadly based. Their major weaknesses are lack of historical validation of weighting schemes and intangible judgments required from umpire personnel.

Systematic development of weighting systems is difficult and complex. NPRDC has utilized Delphi techniques to produce weights for different items based on the opinions of 25 experienced infantry combat commanders (NPRDC, 1975). Their level of focus ranged from squad through brigade. The effort was designed to produce experience-based weighting for use in the Marine Corps Tactical Warfare Simulation, Evaluation, and Analyses System (TWSEAS) discussed below. A similar effort, sponsored by the Defense Advanced Research Project Agency, is currently underway. It utilizes Bayesian decision analysis to produce weightings for items in the Marine Corps Combat Readiness Evaluation System (MCCRES). Both of these efforts have the limitation of not being historically based or validated, although both do use the experience of combat veterans as their main source of insight.

The Marine Corps Tactical Warfare Simulation Evaluation and Analyses System (TWSEAS) merits a brief discussion. Designed for use during an exercise, this automated data processing system provides feedback to units in the field based on actions taken and the situation. Hence, it is possible to reward good performance and punish errors while the unit is in the field. Equally important, the system allows assessment of casualties and can be used to introduce appropriate reductions in unit strength and see how the organization functions. The research team reviewed a plan for Operational Readiness Evaluations developed by the 1st Marine Amphibious force for use in 1976 (1st Marine Amphibious, 1976) as well as Developmental Bulletin 7-76, "Tactical Warfare Simulation, Evaluation and Analysis System" (Marine Corps Development and Education

Command (MCDEC), 1976). As MCDEC indicates, the TWSEAS system provides several significant capabilities. Among the most important are

- The ability to produce scenarios reflecting specific missions and emphasizing key combat functions and different-sized forces,
- The ability to alter weapon availability, threat, and/or other key variables, and
- The generation of a record of events, a set of standard reports reflecting the flow of the exercise and the potential for calculating indices or parameters based on the exercise.

The TWSEAS system can contribute significantly to the suggested system for exercise evaluation.

One final class of evaluation system should be discussed. The most relevant example is MCCRES. This system, currently being developed and implemented, provides a different view of evaluation by structuring the checklists for evaluations. Rather than a simple list of actions to complete, the "mission performance standards" in MCCRES consist of three key elements:

1. Tasks or categories of activities that a unit must carry out (for example, attack with tanks and infantry on converging axes),
2. Conditions that help to define the situation which the unit must overcome (for example, under cover of darkness, or with aggressor forces positioned to oppose both the tanks and the maneuvering infantry), and
3. Requirements or defined standards to help the evaluators in measuring success or failure (for example, fire planning must provide protection for maneuvering tanks).

Given this information, the evaluator needs much less intuition or judgment to complete a "checklist." The task is made even more objective by requiring that the unit be rated as satisfactory or unsatisfactory -- no number or letter grade need be assigned. Either the unit completed the requirements under the specified conditions or it did not.

Of course, MCCRES is a comprehensive system, involving hundreds of tasks, conditions, and requirements for each type of unit. Priorities must be established for types of tasks to be rated as crucial versus noncrucial, tasks to be tested universally versus tasks that should be tested intermittently or only when preparing for specific types of mission, and so forth. The Army has adopted a similar structure for its Army Training and Evaluation Program (ARTEP, 1976).

In summary, U.S. military units in general and U.S. Marine Corps Units in particular have developed increasingly sophisticated systems for evaluating unit performance on exercises. All these systems, however, depend on judgmental data for their structures and weighting, and <u>none</u> of them evaluate the unit's ability to adapt to its field environment. The suggestions made in the next section focus on

- Generating a way of thinking that allows somewhat more objective components in evaluation,
- Using the research results reported earlier in this volume, and other historically validated research, to help establish weightings,
- Taking advantage of the excellent resources already available for evaluation,
- Treating adaptive behavior as a meaningful indicator of unit performance, and
- Developing a data base that will allow establishment of more objective norms for unit performance.

EVALUATION AND EXPERIMENTAL DESIGN

A visit by the principal investigator to the Marine Corps Air-Ground Combat Training Center (MCAGCTC) at Twenty-Nine Palms, California, and observation of field exercises on that base suggested that a key element was being added to the exercise system in the Marine Corps -- the idea of comparability. A series of exercises involving live fire (and therefore no major live aggressor play) has been conducted at Twenty-Nine Palms for battalions from both the East and West Coasts. All battalions have been given roughly the same mission, all faced the same terrain, and, with some variation, and all have encountered the same scenario obstacles.

This element of comparability, particularly in light of the existence of TWSEAS technology, suggests that the Marine Corps is in a unique position to establish an evaluation system for projecting combat effectiveness from exercise data. Such a system would have two principal modes of implementation:

1. To produce a baseline data set, validate the concept, and gain experience in its implementation, a phase in which a single, replicable exercise is evaluated according to experimental design and quasi-experimental design principles.

2. A later phase in which the assumptions of the design are relaxed and data from different exercises, missions, terrain, and other key features are collected and compared with the baseline system.

There is an assumption that repeating similar missions on similar terrain is not a bad thing. MCAGCTC has already made this assumption in choosing to repeat the same problem at Twenty-Nine Palms. The logic is the same as that of the history professor who always gave 10 questions on his final exam and drew them for 20 years from a standard set of 50 he had developed in his first year of teaching. Of course, all the sororities and fraternities on campus had a master list of his exam questions and his course was therefore considered a certain "A." Asked about this, he replied, "Anyone who knows the answers to those 50 questions deserves an A in this course." Similarly, the conduct of a successful combined arms operation across difficult terrain may well be, in itself, a real test of a unit without having to vary the problem from unit to unit.

Since units will experience the Twenty-Nine Palms problem over time, however, new and different elements are constantly entering the problem. For example, new threats emerge, new techniques or weapons may be available to the exercise battalion, levels of support may be different, and the weather will certainly be different for each unit. Rather than classical experimental design (ideal for running rats through mazes), reliance on quasi-experimental design (Campbell and Stanley, 1963) techniques is wise. These techniques call for creating a baseline of data on key variables and then performing both simple statistical associations and more elaborate multivariate techniques to establish "causal relationships."

WHAT DATA CAN BE COLLECTED

Any data from any source can be utilized. Scores from the MCCRES system would make an excellent starting point since MCCRES is both comprehensive and standardized throughout the Marine Corps. However, there is a great deal of highly structured information currently passing through the hands of umpires and the Troop Exercise Control Center (TECC) that could be extremely valuable. For example, interviews at Twenty-Nine Palms indicate that umpires are currently recording, for each target attacked,

- The type and time of intelligence information received, which indicates the presence of a target,
- The time at which the target is acquired by the unit,
- The time at which fire is brought on the target (direct or indirect), and
- The time at which the target is declared neutralized, that is, hit by sufficient firepower to destroy it.

Figure 1 shows the type of data that might be generated from this information. The horizontal axis reflects hours into the exercise. The vertical axis is the time from the identification of a target by the unit

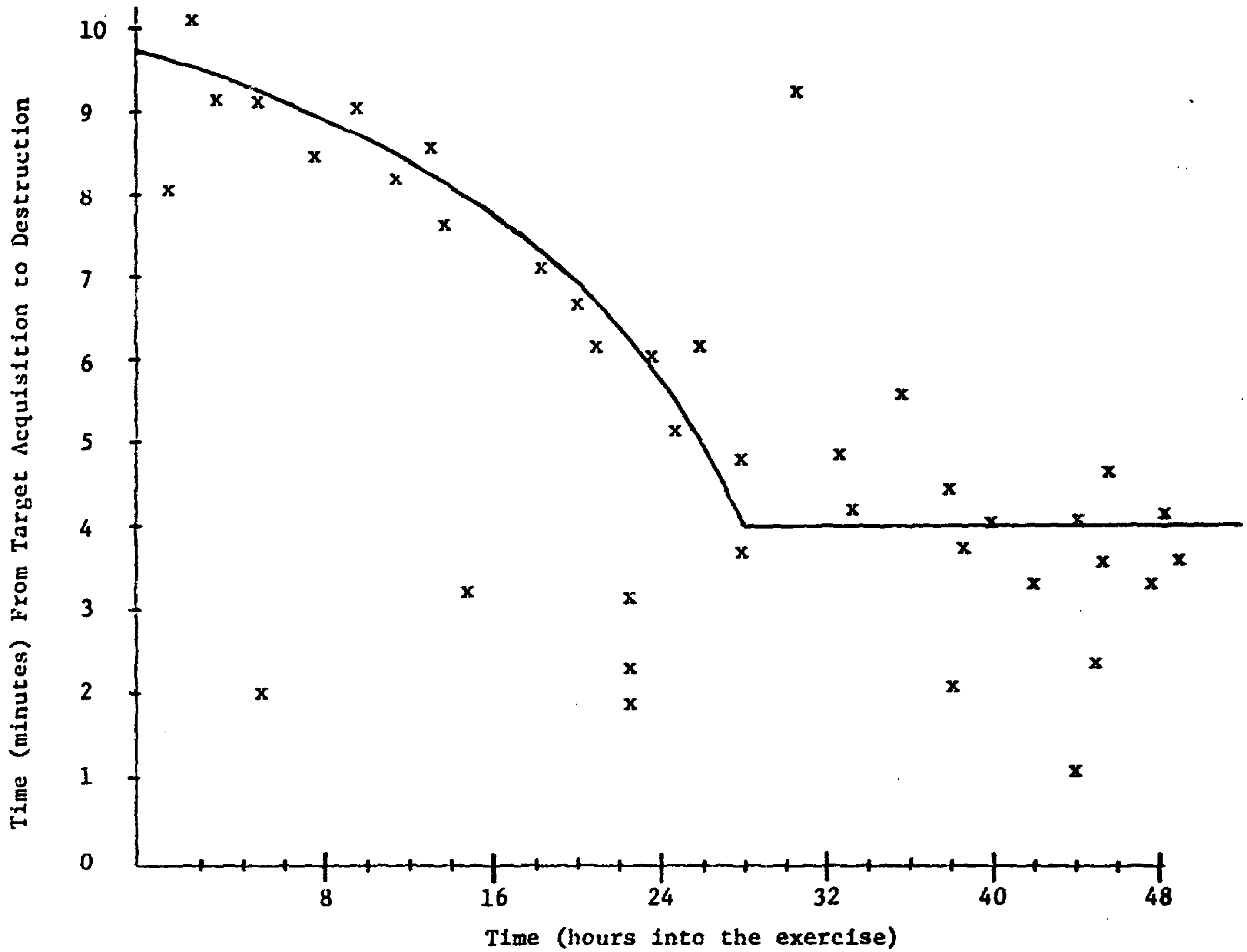

Figure 1. Illustrative Plot of Performance Data

(target acquisition) until the target is declared destroyed by the TECC. Note that these data reflect two things, the relative skill level of the unit at the time the exercise began (needing 8-10 minutes to destroy a target), and the learning or <u>adaptive</u> behavior of the unit. Given the importance of factors like maneuver during action, implementation of the principles of war, use of supporting fires, and so forth, in the research reported earlier, this type of learning behavior is an outstanding indicator of the quality of unit performance.

It is not the purpose here to design a full evaluation system, but a brief list of the types of key indicators that are directly available and could be coded for units would include

- Speed of mission completion (where missions are comparable),
- Speed of response to enemy air threat,
- Variety of weapons utilized,
- Reaction (percentage success and speed) to electronic warfare,
- Speed of destruction of enemy armor, and
- Percentage of time out of communication with supporting arms.

Added onto a system like MCCRES and supported with a system like TWSEAS, this type of data would be invaluable in increasing the objectivity of evaluation and its longer-term utility.

RELATED ANALYSES

Two key types of related analyses would also become possible. First, weighted performance scores could be produced based on the results of this project, the ARPA-sponsored Bayesian efforts, and/or the NPRDC Delphi research. Second, and perhaps more important, research into the linkages between exercise performance measures and other variables would be practical. For example,

- The CACI finding that completion of full unit training cycles is a predictor of combat performance could be examined for exercise situations;
- Unit maintenance ratings could be compared with exercise performance measures;
- Personnel turbulence variables could be used to estimate the impact of rotation policies on performance levels; and
- Number of weather-related injuries might be correlated with unit performance to establish overall levels of preparedness.

These analyses necessarily depend on the gradual development of a large enough data base to perform meaningful statistical analyses. Once sufficient information is present for this baseline data, comparisons with other types of exercises, perhaps already collected through TWSEAS or another system, could be used to produce evaluations. This "system" is, then, a collection of techniques and approaches already in existence but structured for effective analyses and weighted on the basis of research into effective combat performance.

BIBLIOGRAPHY

1st Battalion, 7th Marines (no date) Enclosures to LOI for Tactical Inspection and Test of 1st Bn, 7th Marines.

1st Marine Amphibious (1976) Operational Readiness Evaluations (26 May). Camp Pendleton, CA: USMC, 1st Marine Amphibious.

1st Marine Division (1977) Standard Operation Procedures for Tactical Training Test, Battalion Order P5041.2 (27 June). Camp Pendleton, CA: Headquarters, 1st Marine Division.

1st Marine Division (no date) Evaluation Sheets for 1st Division Order P5041.25.

2nd Marine Division (no date) Appendix C: Unit Training Readiness Indices for Ground Units.

3rd Marine Division (1976a) Predeployment Company Tac Test (28 June). FPO San Francisco: 3rd Marine Division.

_____ (1976b) Company/Platoon Tactical Testing (9 March). FPO San Francisco: 3rd Marine Division.

Army Training and Evaluation Program (ARTEP) (1976) Army Training and Evaluation Program for Light Infantry Battalions (Infantry, Airborne, Air Assault and Ranger). Washington, D.C.: Headquarters, Department of the Army.

CAMPBELL, D. and J. STANLEY (1963) Experimental and Quasi-Experimental Designs for Research. Chicago, IL: Rand McNally and Co.

Department of the Navy, HQ U.S. Marine Corps (5 January 1977) Mission Performance Standards for Marine Corps Combat Readiness Evaluation System (MCCRES). Washington, D.C.: Commandant of the Marine Corps.

_____ (2 March 1977) Marine Corps Combat Readiness Evaluation System (MCCRES) Conference, 15-17 February 1977. Washington, D.C.: Commandant of the Marine Corps.

_____ (28 October 1976) Marine Corps Combat Readiness Evaluation System (MCCRES). Washington, D.C.: Director, Operations Division.

_____ (16 September 1976) Development of Marine Corps Combat Readiness Evaluation System (MCCRES). Washington, D.C.: Director, Operations Division.

Marine Corps Development and Education Command (1976) Tactical Warfare Simulation, Evaluation, and Analysis System (TWSEAS). Quantico, VA: Commanding General, Marine Corps Development and Education Command.

McQUIE, R. (1969) Dynamics of Fire and Maneuver (FIRMA III), Appendix III, Multivariate Analysis of Combat (A Quantitative Analysis). Washington, D.C.: Planning Research Corporation.

_____ (1968) Multivariate Analysis of Combat. Washington, D.C.: Planning Research Corporation.

Mission Performance Standards for Infantry Battalion Missions (Chapter 2) (no date)

Naval Personnel Research and Development Center (NPRDL) (1975) Development of Unit Performance Effectiveness Measures Using Delphi Procedures. San Diego, CA: NPRDC.

_____ (1974) Survey of Unit Performance Effectiveness Measures. San Diego, CA: NPRDC.

PELTIER, L.C. (1964) A Coding System for Describing Factors Related to Military Operations. McLean, VA: Research Analysis Corporation.

U.S. Marine Corps (15 October 1976) Annual Training Plan, Battalion Order 1500.22. FPO San Francisco: U.S. Marine Corps.

_____ (1975) BLT Tactical Evaluation Operation Order (29 March). FPO San Francisco: 3rd Marine Division.

_____ (1974) Marine Infantry Battalion. Publication FMFM 6-3. Washington, D.C.: U.S. Government Printing Office.

_____, Force Troops, Fleet Marine Force, Pacific (15 November 1976) Standing Operating Procedures for the Marine Corps Air-Ground Combat Training Center (MCAGCTC) Troop Exercise Control Center Functions. Twentynine Palms, CA: Commanding General.